


```
Db 121 CGACCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGAGTACGTC 180
Qy 181 GACAGCACACGCGGACAGCTCATGCGCTCTGTGCGGATGGGTCCCGGTGCGTC 240
Db 181 GACAGCACACGCGGACAGCTCATGCGCTCTGTGCGGATGGGTCCCGGTGCGTC 240
Qy 241 AGCAGCCTGACGCTGCGGTGAGTGAACCGCGCGCGAGCTCGGTGCGAGTTCA 300
Db 241 AGCAGCCTGACGCTGCGGTGAGTGAACCGCGCGCGAGCTCGGTGCGAGTTCA 300
Qy 301 CTGCGGCGCGACGCTCAGCTGCGAGGTGCGAGCGAGCCCGACCGCGATCGGT 360
Db 301 CTGCGGCGCGACGCTCAGCTGCGAGGTGCGAGCGAGCCCGACCGCGATCGGT 360
Qy 361 TCCTGTCAACCGCGTGGCCAGCCCGCGCGGAGATCCCGCGATCTGGCAGCCGTA 420
Db 361 TCCTGTCAACCGCGTGGCCAGCCCGCGCGGAGATCCCGCGATCTGGCAGCCGTA 420
Qy 421 GCGCCGTTCTCGTCCGTGACCTTCTGTGCGCTCTCTCTCACTGGAGTTGCGGAGGC 480
Db 421 GCGCCGTTCTCGTCCGTGACCTTCTGTGCGCTCTCTCTCACTGGAGTTGCGGAGGC 480
Qy 481 AGGCAGACACCCAGAAAGGAGAGGGAGCGCGCATCTCGGGACCGGGAAACCGGAG 540
Db 481 AGGCAGACACCCAGAAAGGAGAGGGAGCGCGCATCTCGGGACCGGGAAACCGGAG 540
Qy 541 CCGAGAGGTGGCCCGCAGGAAAGCGTCCCGCAGCGCGAAGCGAGGGCGATGTT 600
Db 541 CCGAGAGGTGGCCCGCAGGAAAGCGTCCCGCAGCGCGAAGCGAGGGCGATGTT 600
Qy 601 CAGGCTGAGGAGCAGCGCGGTGGGAGCAGCACCACCGAGGACCGGAAACAGTGTG 660
Db 601 CAGGCTGAGGAGCAGCGCGGTGGGAGCAGCACCACCGAGGACCGGAAACAGTGTG 660
Qy 661 GCGGCGATGCGAGCTGAGCTGGGAGCGAGGAGCCATGTCTCGCGAGAGTGGCTCAG 720
Db 661 GCGGCGATGCGAGCTGAGCTGGGAGCGAGGAGCCATGTCTCGCGAGAGTGGCTCAG 720
Qy 721 GACCCCGTGATTCGGATCCCTCGAGACGAGGCTCCCGCAGCAAGCAGCGAGCAGAC 780
Db 721 GACCCCGTGATTCGGATCCCTCGAGACGAGGCTCCCGCAGCAAGCAGCGAGCAGAC 780
Qy 781 AAGGAGCGCTGCGTTCCAGTCTTAGAGCAGAAAGTACGGCTACTATCTGCAAGGAC 840
Db 781 AAGGAGCGCTGCGTTCCAGTCTTAGAGCAGAAAGTACGGCTACTATCTGCAAGGAC 840
Qy 841 TGCAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGCGAGGACCAAGTAAGTGTTA 900
Db 841 TGCAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGCGAGGACCAAGTAAGTGTTA 900
Qy 901 CTTCAACAGTTCTGCCGAGTGTGAGAAATCTTACAACTTACAGAGTGGAGACAT 960
Db 901 CTTCAACAGTTCTGCCGAGTGTGAGAAATCTTACAACTTACAGAGTGGAGACAT 960
Qy 961 CACTGTCAAAAGTTGTAAGAACTAGATGTGCTGCCAGTCAAGTTTCGCCAGTGGTA 1020
Db 961 CACTGTCAAAAGTTGTAAGAACTAGATGTGCTGCCAGTCAAGTTTCGCCAGTGGTA 1020
Qy 1021 CCCTAAACGCCCCATCGGCAAGACTGTGTGGGAGATGCAAGCAAAACGCTGTCTTG 1080
Db 1021 CCCTAAACGCCCCATCGGCAAGACTGTGTGGGAGATGCAAGCAAAACGCTGTCTTG 1080
Qy 1081 CGACGACCTTCACTTCAATAATCATATTTAGTGGAGTGCAGAAAGTTCCTCTAGA 1140
Db 1081 CGACGACCTTCACTTCAATAATCATATTTAGTGGAGTGCAGAAAGTTCCTCTAGA 1140
Qy 1141 TGGGGCTAATGGAATGGAAGTGGAGCTTTCTCCCTCTTCACTCTTCCCTTTCCAAAT 1200
Db 1141 TGGGGCTAATGGAATGGAAGTGGAGCTTTCTCCCTCTTCACTCTTCCCTTTCCAAAT 1200
Qy 1201 TCCTTACACAGACAGTGTACTTGGATATAAGCTGTGAATAAAGGTATTGCAACA 1260
Db 1201 TCCTTACACAGACAGTGTACTTGGATATAAGCTGTGAATAAAGGTATTGCAACA 1260
```

```
Db 1201 TCCTTATGACAGACAGTGTACTTGGATATAAAGCTGTGAATAAAGGTATTGCAACA 1260
Qy 1261 AAAAAAAAAAAAAAAAAA 1277
Db 1261 AAAAAAAAAAAAAAAAAA 1277
RESULT 2
US-09-844-864-18
; Sequence 18, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 0907797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 814
; TYPE: DNA
; ORGANISM: mus musculus
; US-09-844-864-18
Query Match 62.6%; Score 799.4; DB 9; Length 814;
Best Local Similarity 99.9%; Pred. No. 1.4e-228;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 GCGGCGAGCGCGGAGCAGCACCATGTTCCCGCGAGCAGCTTCCACCCCTGCCCGCA 62
Db 1 GCGGCGAGCGCGGAGCAGCACCATGTTCCCGCGAGCAGCTTCCACCCCTGCCCGCA 60
Qy 63 TCCATTATCCGAGCGCACCAAGCGGGATGGGTGGAGGTTCCGAGCCAGGGGTGCGG 122
Db 61 TCCATTATCCGAGCGCACCAAGCGGGATGGGTGGAGGTTCCGAGCCAGGGGTGCGG 120
Qy 123 ACCGCGCCCCCTCTCTTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTGCA 182
Db 121 ACCGCGCCCCCTCTCTTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTGCA 180
Qy 183 CAGCCACAGCGGCGACAGCTCATGCGGCTCTGTGCGGATGGGTCCCGGTGCGTCA 242
Db 181 CAGCCACAGCGGCGACAGCTCATGCGGCTCTGTGCGGATGGGTCCCGGTGCGTCA 240
Qy 243 CAGCGTGAGCGCTGCGGTGACAGTGAACCCCGCGCGACGCTCGGTGCAAGTGTCACT 302
Db 241 CAGCGTGAGCGCTGCGGTGACAGTGAACCCCGCGCGACGCTCGGTGCAAGTGTCACT 300
Qy 303 CCGGCGCGACGCTGCAAGCTGCGAGGTGCGAGCCAGCCCGACGCCCGATCGGGTTC 362
Db 301 CCGGCGCGACGCTGCAAGCTGCGAGGTGCGAGCCAGCCCGACGCCCGATCGGGTTC 360
Qy 363 CTGTCAACCCCGTGGCCAGCGCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGC 422
Db 361 CTGTCAACCCCGTGGCCAGCGCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGC 420
Qy 423 CCGGTTCGTGCTGCGTACCTTCTGTGGCTCTCTCTCTCTCACTGGAGGTTGCGGAGGAG 482
Db 421 CCGGTTCGTGCTGCGTACCTTCTGTGGCTCTCTCTCTCTCACTGGAGGTTGCGGAGGAG 480
Qy 483 GCAGACACCCAGAAAGGAGAGGGAGCCCGCATCTTCGGGGAACCCCGAGGCC 542
Db 481 GCAGACACCCAGAAAGGAGAGGGAGCCCGCATCTTCGGGGAACCCCGAGGCC 540
Qy 543 GAGAGAGTGGCCCGGAGAAAGCGGTCCCGCAGCGAGGAGGAGGCGATGTCA 602
Db 543 GAGAGAGTGGCCCGGAGAAAGCGGTCCCGCAGCGAGGAGGAGGCGATGTCA 602
```



```

RESULT 8
US-09-844-864-20
; Sequence 20, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Xueheng
; APPLICANT: Wu, Xiongwei
; TITLE OF INVENTION: OVARY SPECIFIC C
; FILE REFERENCE: P019250S2 / 09807799
; CURRENT APPLICATION NUMBER: US/09/844
; CURRENT FILING DATE: 2001-04-27
; PRIORITY APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/2
; PRIOR FILING DATE: 1999-10-28

```


; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 105
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-20

Query Match 5.6%; Score 72; DB 9; Length 105;
Best Local Similarity 100.0%; Pred. No. 3.9e-11;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 899 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 958
Db 19 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 78

QY 959 ATCACCTGTCAA 970
Db 79 ATCACCTGTCAA 90

RESULT 9

US-09-844-864-24
; Sequence 24, Application US/09844864
; Patent No. US20020042926A1

; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin

; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei

; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48

; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28

; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 24
; LENGTH: 105
; TYPE: DNA

; ORGANISM: mus musculus
US-09-844-864-24

Query Match 5.5%; Score 70.4; DB 9; Length 105;
Best Local Similarity 98.6%; Pred. No. 1.2e-10;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 899 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 958
Db 19 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 78

QY 959 ATCACCTGTCAA 970
Db 79 ATCACCTGTCAA 90

RESULT 10

US-10-437-963-8597/c

; Sequence 8597, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 8597
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Oryza sativa

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15091C.1
US-10-437-963-8597

Query Match 4.3%; Score 55.2; DB 17; Length 1041;
Best Local Similarity 46.2%; Pred. No. 1.3e-05;
Matches 258; Conservative 0; Mismatches 293; Indels 7; Gaps 2;

QY 89 GGGATGGCTGGAGTTCCGAGCCAGGGGCTGCGAGCCCGCGCCCCCTCTCTCTCTCCCCG 148
Db 583 GGGGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 524

QY 149 GCTACAGACAGCTCATGCGCGCGGGAGTACGTGACAGCACCAGCGGGGACAGCTCATGG 208
Db 523 GCGGGGCGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGG 465

QY 209 CCCTGCTGTGCGGGATGGGTCCCGGCTCGGTACAGACCGTACGCTGCGGTGACGTGA 268
Db 464 GCGCGGGGCGGGGCGGGTGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGG 405

QY 269 ACCCGCGCGCGGACG-----CCTCGGTGAGTGTTCCTCGGGCGCGCGCGCTGCAGC 322
Db 404 CGCGCGGTGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345

QY 323 CTGCAGGTGCGCGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 382
Db 344 GGGGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285

QY 383 CCGGCGCGGGAGATCCCGCGGATCTTGGCAGACCGTAGCCCGTTCTCTCTCTGACCT 442
Db 284 GGTGCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 225

QY 443 TCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 502
Db 224 GGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 165

QY 503 AGGGAGCGCGGATCTCTCGGGACCGGGAAACCGAGAGCGCGAGAGGTGGCGCGGAGGA 562
Db 164 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105

QY 563 AAGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 622
Db 104 CCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 45

QY 623 GGGAGCAGCAGCCAC 640
Db 44 GCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27

RESULT 11

US-10-412-699B-1133

; Sequence 1133, Application US/10412699B

; Publication No. US20040045049A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Zhang, James

; APPLICANT: Fromm, Michael E.

; APPLICANT: Heard, Jacqueline E.

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Adam, Luc J.

; APPLICANT: Broun, Pierre E.

; APPLICANT: Pineda, Omaira

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James S.

; APPLICANT: Yu, Guo-Liang

```

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1133
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-412-699B-1133

Query Match 4.2%; Score 54.2; DB 13; Length 1500;
Best Local Similarity 43.4%; Pred. No. 3.1e-05;
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

Qy 137 CCTTCTCCCGGCTACAGACGCTCATGGCGGAGTGGTTCCTGTCGATCGGTCGACGCGTG 196
Db 298 CTTCTTCCACCCCGGAGGCTCTCGAAGAGGAGTACTACGACGAGCTCCCGG 357

Qy 197 CACAGCTCATGGCCCTGCTGCGGGATGGGTCCCGGTGCGTACGACGCGGTGACGCTG 256
Db 358 AGAAGAAGCGGCGCTCACCGCGGAGGAGTGCATCTGCTGAGAGGAGCTTCGAGGAGG 417

Qy 257 CGGTGAGTGAAACCGGCGCGGAGCGCTCGTGTGAGTGTCTACTGGGCGCGCAGCG 316
Db 418 AGAACAAGCTGGAGCGGAGGAGGAGCGAGCTGGCGGGAAGCTAGGCTGACGCGCG 477

Qy 317 TGCAGCTGCGAGGTGCGGAGCGAGCGCGGATCGGTTCTGTGTCACCCCGTG 376
Db 478 GGCAGGTGCGGTGTGTTCCAGAACCGCGCGCGCTGGAAGACCAAGCAGCTCGAGC 537

Qy 377 GCCAGCGCGGCGCGGAGATCCCGCGGATCTGCGACAGCGTACGCCCGCTTCGTCCG 436
Db 538 GCGACTTCGACCGCTCAAGGCGTGTTCGAGCGCCCTCCGCGCGCACCGCCCTCC 597

Qy 437 TGACCTTCTGCGCTCTCTCTCTACTGAGGTGTCGGGAGGAGGAGGAGCAGACCCAGA 496
Db 598 TCCAGGACAAACCGCGCTCCACTCTCAGTCTAGTCTGTTGACCGAGAGCTGCAAGAGA 657

Qy 497 AGGAGAGGAGGAGCGCGGATCTCTCGGAGACCGGGAACCGGAGCGGAGAGGTGGCGG 556
Db 658 AGGAGACGACACCGAGGAGGAGCGCGGCGGCGGTGACGTCCCGGCGCTGCTGCGG 717

Qy 557 CGAGAAAGCGGTCCCCCAGCGCGGAGGAGGCGGATGTTTTCAGGCTGACGAGGCGG 616
Db 718 CGGCGGAGTGAAGGTCCCGGTCCCGGAGCGCGGAGGAGGAGGAGGCGGCGG 777

Qy 617 CCGGTGGAGCAGCAGCCACCGGAGGAGCCGGAACAGTGTGGCGCGGATGACGCTG 676
Db 778 CGGCGTTGAGGAGCAGCAGGAGGAGGAGGTGAAGGCGGAGGAGGCTGAGCACGGCA 837

Qy 677 AGCTGGGAGGAGGAGGAGGATGCTCTCGCGCAGAGATGG 715
Db 838 GCGGCGGAGCGCGGTGTGTGACACGAGGAGCGCAACTGG 876

RESULT 12
US-10-374-780A-680
; Sequence 680, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MEI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 680
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G409
US-10-374-780A-680

Query Match 4.2%; Score 54.2; DB 16; Length 1500;
Best Local Similarity 43.4%; Pred. No. 3.1e-05;
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

Qy 137 CCTTCTCCCGGCTACAGACGCTCATGGCGGAGTGGTTCGACGAGTACGACGAGCGG 196
Db 298 CTTCTTCCACCCCGGAGGAGTCTCGAAGAGGAGTACTACGACGAGCTCCCGG 357
```

QY 197 CACAGCTCATGGCCCTGCTGTGCGGATG3GTTCCCGGTGCGTGCAGACGCGTGAAGCTG 256
Db 358 AGAAGAGCGGCGCTTCAAGCGGAGAGAGTGCATCTGCTGGAGAGAGTTTCAGAGAG 417
QY 257 CGGTGAGGTGAACCCGCGCGCGAGCGCTTCTGCTGAGTGTTCAGTGGGCGCGCAGCG 316
Db 418 AGAACAAGCTGGAGCGGAGCGGAGAGAGAGTGGCGCGGAAAGTAGGCTGAGCGCG 477
QY 317 TGCAGCTTCAGAGGTGCGGAGCGAGCGCGCGAGCGCGCGATCGGTTCTGTCAACCCGCTG 376
Db 478 GCGAGTTCGCGCTGTGTTCCAGAACCGCGCGCGCTGGAAGACCAAGAGCTGAGC 537
QY 377 GCCACGCGCGCGCGGAGATCCCGGATCTCTGGAGACCGTAGCGCCCTTCTGTCTCGTCG 436
Db 538 GCGACTTTCGACCGCTCAAGCGTGTTCAGAGCTTCCGCGCGAGCACGAGCGCTCTC 597
QY 437 TGACCTTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 496
Db 598 TCCAGGACAAACACCGCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657
QY 497 AGGAGAGGGGAGCGCGGATCTCTCGGAGACCGGAAACCGGAGCGGAGAGGTGGCG 556
Db 658 AGGAGAGCGAGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
QY 557 CGAGGAAAGCGTCCCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 616
Db 718 CGGCGAGCTGAAGTTCGCGCTCCGAGCGCGGAGGAGCGGAGCGGAGCGGAGCGGAG 777
QY 617 CGGCGTGGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 676
Db 778 CGGCGTTCGAGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 837
QY 677 AGCCTGCGGAGCGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 715
Db 838 GCGGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 876

RESULT 13

US-10-437-963-52312/c
; Sequence 52312, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 52312
; LENGTH: 2664
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(2664)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54620C.1
US-10-437-963-52312

Query Match 4.1%; Score 52.8; DB 17; Length 2664;
Best Local Similarity 49.0%; Pred. No. 0.00011;
Matches 141; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 172 GAGTACGTGCAGACGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 231

Db 1555 GAGAGAGGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1496
QY 232 CGGTGCGTGCAGAGCGGTGAGCGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 291
Db 1495 CGCGCGCGCGTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1436
QY 292 CAGTGTTCAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 351
Db 1435 CCGCGCATGTGCGCTGCTCCCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1376
QY 352 CGATCGGCTTCTGTCAACCCCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 411
Db 1375 GCGCGCGTCCCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1316
QY 412 CAGACGTGAGCGCGCGTTCGCTCGCTGAGCGCTTCTGTGCGCTCTCTCTCTCTCTCT 459
Db 1315 CTGTCCCGCGCGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1268

RESULT 14

US-10-437-963-77205/c
; Sequence 77205, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77205
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(720)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77125C.1
US-10-437-963-77205

Query Match 4.1%; Score 52.2; DB 17; Length 720;
Best Local Similarity 48.2%; Pred. No. 8.6e-05;
Matches 176; Conservative 0; Mismatches 188; Indels 1; Gaps 1;
QY 308 GCGCAGCGTGCAGCGTGCAGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
Db 605 GCGCGCACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY 368 AACCCCGTGGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 427
Db 545 GAGCGGTGCGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487
QY 428 TCTGCTCGGTGAGCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487
Db 486 TGGGTCGAGAGCGTGTGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 427
QY 488 CACCCACGAGGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 547
Db 426 CACCCCTGCGAGTGGCGTGCAGGCGTGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
QY 548 AGTGTGCGCGAGGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 607

Db 366 AGTGGCGCGCGCCACCGGAGCAGAGGAGGAGGAGCGGAGCGCGGAGCGCGGAGGCG 307
Qy 608 CAGGCGAGGCGCGGTGGAGCAGCAGCCACACCGAGGACCGGAACAGTGTGCGGCGCA 667
Db 306 CCGCGCGGAGGCGCGGTGGAGAGCGCGCGCGGTGGAGCGCCAGCAGCGCGCGGTGGG 247
Qy 668 TGCAG 672
Db 246 AGCG 242

RESULT 15

US-10-084-846A-99
; Sequence 99, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREPFER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 165.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 99
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: aviG4 dna: partial sequence of coding strand 1; nucleotide 1
; OTHER INFORMATION: corresponds to nucleotide 45,341 of coding strand 1.
US-10-084-846A-99

Query Match 4.0%; Score 51.6; DB 16; Length 987;
Best Local Similarity 49.6%; Fred. No. 0.00015;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
Qy 80 CCAAAGCCGGGATGGCTGGAGGTTCCGGAGCCAGGGGCTGCCGCCGCCGCCGCCCTCCT 139
Db 260 CCGCGTCCGACGCGCTGGCAGCACCTGGACACAGCGTCCGACCCGGCGCCCGGCT 319
Qy 140 TCCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTGACAGCCACCGAGCGGCAC 199
Db 320 TCGACCAAGTCTCGGCGGGACTTCTCGGCTACTCGGCGGACGACCCGACCTGCTGT 379
Qy 200 AGCTCATGGCCCTGCTGCGGGATGGTCCCGGTCGCTCAGCAGCGGTGAGCGGTGGG 259
Db 380 GGCTCTACACGGCGGATGAGCCAGGACACCGTGGGATCGCGGCTGTGCGCGGC 439
Qy 260 TGCAGGTGAACCCGCGCGCGCCTCGGTGTCAGTGTTCACCTCGGCGCGCGCACGCTGC 319
Db 440 ACCAGGACTTCTCGGCGTCCGACCGGTGGTGGAGCTCGGCGCGCGGACGCGGCTGC 499
Qy 320 AGCTGAGGTTGCGAGCCAGCC 345
Db 500 TGGCGCGGTGCTGCGCGCGACCC 525

Search completed: September 24, 2004, 20:14:47
Job time : 680 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 17:12:32 ; Search time 3653 Seconds
(without alignments)
10439.092 Million cell updates/sec

Title: US-09-830-810A-1
Perfect score: 1277
Sequence: 1 aagcgggcgagggcgcgga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estm:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_hc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	643	50.4	675	10	BB641267
2	642.4	50.3	666	10	BB264222
3	591.4	46.3	615	10	BG071693
4	582.4	45.6	595	12	BG084538

C	5	570.4	44.7	594	10	BF471866	BF471866 UI-M-BH3-
C	6	497.2	38.9	525	10	BB704019	BB704019 BB704019
C	7	473.8	37.1	499	9	AI854700	AI854700 UI-M-BH0-
C	8	448.4	35.1	491	10	BB703259	BB703259 BB703259
C	9	413.8	32.4	491	10	BB704648	BB704648 BB704648
C	10	405.8	31.8	436	9	AU044294	AU044294 AU044294
C	11	398.2	31.2	425	10	BB705931	BB705931 BB705931
C	12	396.4	31.0	590	9	AU023153	AU023153 AU023153
C	13	393	30.8	450	10	BE946858	BE946858 UI-M-BH3-
C	14	392.2	30.7	419	10	BB704449	BB704449 BB704449
C	15	367.4	28.8	521	10	BB703869	BB703869 BB703869
C	16	367	28.7	427	10	BB706957	BB706957 BB706957
C	17	364.2	28.5	434	10	BB693732	BB693732 BB693732
C	18	347.2	27.2	391	10	BB704146	BB704146 BB704146
C	19	333.8	26.1	400	10	BB700620	BB700620 BB700620
C	20	319	25.0	905	28	B2114495	B2114495 CH230-421
C	21	316.2	24.8	441	10	BB704848	BB704848 BB704848
C	22	282.2	22.1	301	13	EX637052	EX637052 EX637052
C	23	264.6	20.7	315	10	BB248342	BB248342 BB248342
C	24	227.4	17.8	642	13	BY719704	BY719704 BY719704
C	25	224.8	17.6	654	28	AZ225053	AZ225053 RPCI-23-6
C	26	211.6	16.6	251	9	AV359580	AV359580 AV359580
C	27	208.6	16.3	251	9	AV358045	AV358045 AV358045
C	28	203.2	15.9	249	9	AV357766	AV357766 AV357766
C	29	203	15.9	290	10	BB063113	BB063113 BB063113
C	30	201.8	15.8	539	9	AL904756	AL904756 AL904756
C	31	200.8	15.7	909	13	EX755136	EX755136 BX755136
C	32	200	15.7	286	10	BB060381	BB060381 BB060381
C	33	199.8	15.6	246	10	BB060134	BB060134 BB060134
C	34	198	15.5	539	9	AL904892	AL904892 AL904892
C	35	197.8	15.5	642	12	BM861479	BM861479 fy46d03.x
C	36	197	15.4	643	12	BU098465	BU098465 BU098465
C	37	196.8	15.4	526	9	AL904839	AL904839 AL904839
C	38	196.8	15.4	534	9	AL904820	AL904820 AL904820
C	39	196.8	15.4	559	9	AL904774	AL904774 AL904774
C	40	196.8	15.4	575	9	AL904855	AL904855 AL904855
C	41	196.4	15.4	576	9	AL904863	AL904863 AL904863
C	42	196.4	15.4	578	9	AL904817	AL904817 AL904817
C	43	196.4	15.4	579	9	AL904815	AL904815 AL904815
C	44	196.4	15.4	579	9	AL904816	AL904816 AL904816
C	45	196.4	15.4	604	9	AL904765	AL904765 AL904765

ALIGNMENTS

RESULT 1
BB641267
LOCUS BB641267 675 bp mRNA linear EST 26-OCT-2001
DEFINITION BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus
musculus cDNA clone A830014H23 5', mRNA sequence.
BB641267
VERSION BB641267.1 GI:16476392
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 675)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,H., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

Location/Qualifiers

1. 666

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A830014H23"

/tissue_type="cortex"

/dev_stage="10 days neonate"

/lab_host="DHL0B"

/clone_lib="RIKEN full-length enriched, 10 days neonate cortex"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCTCTTTTATTAATTAATCCCTCCCGGAGG 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCTTTTATTAATTAATCCCTCCCGGAGG 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1."

ORIGIN

Query Match 50.3%; Score 642.4; DB 10; Length 666;
 Best Local Similarity 99.6%; Pred. No. 6:2e-87;
 Matches 665; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

588 GGAGGGCGATGTTTACGCTCAGCGCAGGCGGGTGGAGCAGCCACCGGAGGA 647
 1 GGAGGGCGATGTTTACGCTCAGCGCAGGCGGGTGGAGCAGCCACCGGAGGA 60

648 CCGGACAGTGTGGCGCGATGAGTCTGAGCTGGAGCGAGGCGATGCTCCGCG 707
 61 CCGGACAGTGTGGCGCGATGAGTCTGAGCTGGAGCGAGGCGATGCTCCGCG 120

708 AGAGATGGCTCAGACCCCGGCTATTCGGATGCCCTCGAGACAGGCTCCCGGAAG 767
 121 AGAGATGGCTCAGACCCCGGCTATTCGGATGCCCTCGAGACAGGCTCCCGGAAG 180

QY 768 CACGAGCAGGACAGGAGCGGCTCGGTTTCAGTTCTTTAGAGCAGAACTACGGCTACTA 827
 Db 181 CACGAGCAGGACAGGAGCGGCTCGGTTTCAGTTCTTTAGAGCAGAACTACGGCTACTA 240

QY 828 TCACCTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCCTATGTTGGTGTGTCAGGGCAC 887
 Db 241 TCACCTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCCTATGTTGGTGTGTCAGGGCAC 300

QY 888 CAGTAAGGTGTTTACTTCAACAGAGTTCTCCGAGTGTGTGAGAAATCCTACACCCCTTACA 947
 Db 301 CAGTAAGGTG-TACTTCAACAGAGTTCTCCGAGTGTGTGAGAAATCCTACACCCCTTACA 359

QY 948 GAGTGGAGCAGATCACCCTGTCAAGTTCTAAAGAACTAGATGTGCTCCCGAGTCAGAT 1007
 Db 360 GAGTGGAGCAGATCACCCTGTCAAGTTCTAAAGAACTAGATGTGCTCCCGAGTCAGAT 419

QY 1008 TTGCGCAGTGGAGCCCTTAAAGCGCCCACTCGGCAAGACTTGTGTGGAGATGCAAGGACA 1067
 Db 420 TTGCGCAGTGGAGCCCTTAAAGCGCCCACTCGGCAAGACTTGTGTGGAGATGCAAGGACA 479

QY 1068 AACGCTGTCTCGGAGCAGCACCCTTCAAGTTCATATCATCATTTAGTGGAGTGCAGAAA 1127
 Db 480 AACGCTGTCTCGGAGCAGCACCCTTCAAGTTCATATCATCATTTAGTGGAGTGCAGAAA 539

QY 1128 CGTTTCTGCTAGATGGGCTTAATGGAATGGAATGAGTGTTCCTCCCTCTTCACTCTCT 1187
 Db 540 CGTTTCTGCTAGATGGGCTTAATGGAATGGAATGAGTGTTCCTCCCTCTTCACTCTCT 599

QY 1188 TCCCTTTCCAAATTTCTTCATGACAGACAGTGTCTTACTTGGATATAAGGCTCTGTAATAAA 1247
 Db 600 TCCCTTTCCAAATTTCTTCATGACAGACAGTGTCTTACTTGGATATAAGGCTCTGTAATAAA 658

QY 1248 GGTATTGC 1255
 Db 659 GGTATTGC 666

RESULT 3
 BG071693/3
 LOCUS H3102B03-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 DEFINITION H3102B03 3', mRNA sequence.
 ACCESSION BG071693
 VERSION BG071693.2 GI:40072037
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 615)
 Tanaka, T., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagara, R., Doi, H., Wood, W.H., III, Becker, K.G. and Ko, M.S.H.
 Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
 Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
 20381348
 10922068
 On Jan 26, 2001 this sequence version replaced gi:12554262.
 Other ESTs: H3102B03-5
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please visit <http://igsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3102 row: B column: 03
 Seq primer: -21M13 Forward
 High quality sequence stop: 615
 POLYA=Yes.

FEATURES Location/Qualifiers
 source 1. 615


```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-3"
/db_xref="taxon:10090"
/clones="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

Query Match 46.3%; Score 591.4; DB 10; Length 615;
Best Local Similarity 99.5%; Pred. No. 2.8e-79;
Matches 614; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 643 GAGGACCGGAACAGTGTGCGGGATGTCAGTCTGAGCCTGGAGCGAGGACGATGCTCT 702
DB 615 GAGGACCGGAACAGTGTGCGGGATGTCAGTCTGAGCCTGGAGCGAGGACGATGCTCT 556
QY 703 GCCGACGAGATGCTCAGGACCGCGGTGATTCGGATGCGCCCTCGAGACCGAGGCTCCCG 762
DB 555 GCCGACGAGATGCTCAGGACCGCGGTGATTCGGATGCGCCCTCGAGACCGAGGCTCCCG 496
QY 763 CAAGCAGCGGACGACAGGAGCGCTGCTTCCAGTCTTAGCAGGAGATGACGCG 822
DB 495 CAAAGCAGCGGACGACAGGAGCGCTGCTTCCAGTCTTAGCAGGAGATGACGCG 436
QY 823 TACTATCACTGCAAGGACTGCAAAATCCGGTGGAGAGCGCTTATGTGTGTGTGCGAG 882
DB 435 TACTATCACTGCAAGGACTGCAAAATCCGGTGGAGAGCGCTTATGTGTGTGTGCGAG 376
QY 883 GGCACCAAGTAAAGTGTACTTCAAAAGTGTCCGAGTGTGTGAGAAATCTTACAAACC 942
DB 375 GGCACCAAGTAAAGTGTACTTCAAAAGTGTCCGAGTGTGTGAGAAATCTTACAAACC 317
QY 943 TTACAGAGTGGAGATCACTGCTCAAGTCTTAAAGAACTAGATGTGCTGCCAGT 1002
DB 316 TTACAGAGTGGAGATCACTGCTCAAGTCTTAAAGAACTAGATGTGCTGCCAGT 257
QY 1003 CAGATTTCGCCAGTGGACCCCTAAACGCCGCCCTCGCAAGACTTGTGTGGAGATGCAA 1062
DB 256 CAGATTTCGCCAGTGGACCCCTAAACGCCGCCCTCGCAAGACTTGTGTGGAGATGCAA 197
QY 1063 GGCACCAAGCTGTCTGCGACGACGACCTTCAAGTCAAAATACATCATTTAGTGAGATC 1122
DB 196 GGCACCAAGCTGTCTGCGACGACGACCTTCAAGTCAAAATACATCATTTAGTGAGATC 137
QY 1123 GAAACGTTTCTGTAGATGGGCTAATGAATCGCAAGTGTCTTCTCCCTCTTCA 1182
DB 136 GAAACGTTTCTGTAGATGGGCTAATGAATCGCAAGTGTCTTCTCCCTCTTCA 77
QY 1183 CCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTGTACTTGGATATAAAGCGCTGTGAA 1242

```

ORIGIN

```

Db 76 CCTCTCCCTTTCCAAATTTCTTCATGACAGACGTG-TACTTGGATATAAAGCGCTGTGAA 18
QY 1243 TAAAGGATTTGCAAC 1259
Db 17 TAAAGGATTTGCAAC 1

RESULT 4
LOCUS BG084538
DEFINITION H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION BG084538
VERSION BG084538.2 GI:40072038
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 595)
AUTHORS Tanaka,T.S., Jazadat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagataja,R., Doi,H.,
Wood,W.H. III, Becker,K.G., and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and
JOURNAL embryo using a 15,000 mouse developmental cDNA microarray
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
PUBMED 20381348
COMMENT On Jan 26, 2001 this sequence version replaced gi:12567102.
Other ESTs: H3102B03-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cdna/15k.html for details.
Plate: H3102 row: B column: 03
Seq primer: -21M13 Reverse
High quality sequence stop: 595
POLYA=No. Location/Qualifiers
1..595
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-5"
/db_xref="taxon:10090"
/clone="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on

```



```

915 GCCGAGTGTGTGAGAAATCTTACAAACCCCTTACAGAGTGGAGGACATCACCTGTCAAAGTT 974
Df      |||
235 GCCGAGTGTGTGAGAAATCTTACAAACCCCTTACAGAGTGGAGGACATCACCTGTCAAAGTT 176
Qf      |||
975 GTAAAGAACTAGATGTGCTGCCAGTCAGATTTGCCAGTGGACCTAAACGCCGCC 1034
Df      |||
175 GTAAAGAACTAGATGTGCTGCCAGTCAGATTTGCCAGTGGACCTAAACGCCGCC 116
Qf      |||
1035 ATCGCGAAGACTGTGTGGGAGATGCAAGGCAACAGCGCTCTCTCGGACAGCACCTTCA 1094
Df      |||
115 ATCGCGAAGACTGTGTGGGAGATGCAAGGCAACAGCGCTCTCTCGGACAGCACCTTCA 56
Qf      |||
1095 GCTTCAATATCATCTTTAGTGGAGTGGAAACGTTTCTCTAGATGGGCTAAT 1150
Df      |||
55 GCTTCAATATCATCTTTAGTGGAGTGGAAAC-TTTCTCTAGATGGGCTAAT 1

RESULT 6
BB704019
LOCUS      BB704019 RIKEN full-length enriched, in vitro fertilized eggs Mus
DEFINITION musculus cDNA clone 7420459B08 3', mRNA sequence.
ACCESSION BB704019
VERSION    BB704019.1 GI:16052854
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus (house mouse)

REFERENCE
AUTHORS    Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Saeki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,P., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
TITLE      Unpublished (2001)
JOURNAL    Contact: Yoshihide Hayashizaki
COMMENT    Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers

```

source

```

1. .525
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420459B08"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/lab_host="DH10B"
/dev_stage="egg"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAAAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCGAGTTAAATAATTAATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"
Query Match 38.9%; Score 497.2; DB 10; Length 525;
Best Local Similarity 99.0%; Pred. NO. 4.1e-65;
Matches 521; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Qf 732 TTCGATGCCCTCGAGACCGAGCTCCCGCAAGACGCGAGGACGAGGCGCT 791
Df 1 TTCGATGCCCTCGAGACCGAGCTCCCGCAAGACGCGAGGACGAGGCGCT 60
Qf 792 GCGTTTCAGTCTTTAGAGCAGAGTACGGCTACTATCACTGCAAGGACTGCAAAATCG 851
Df 61 GCGTTTCAGTCTTTAGAGCAGAGTACGGCTACTATCACTGCAAGGACTGCAAAATCG 120
Qf 852 GTGGAGAGCGCTATGTGTGTGTGCGAGGACCCAGTAAGTGTGTCTTCAACAGT 911
Df 121 GTGGAGAGCGCTATGTGTGTGTGCGAGGACCCAGTAAGTGTGTCTTCAACAGT 179
Qf 912 TCTCCGAGTGTGTGAGAAATCTTACAAACCTTTACAGAGTGGAGGACATCACTGTCAA 971
Df 180 TCTCCGAGTGTGTGAGAAATCTTACAAACCTTTACAGAGTGGAGGACATCACTGTCAA 239
Qf 972 GTTGTAAAGAACTAGATGTGCTGCCAGTGCAGATTTGCCACGCTGGACCTTAAACGCC 1031
Df 240 GTTGTAAAGAACTAGATGTGCTGCCAGTGCAGATTTGCCACGCTGGACCTTAAACGCC 299
Qf 1032 CCCATCGGCAAGCTGTGTGGGAGATGCAAGGCAACAGCGCTCTCTCGGACAGCACT 1091
Df 300 CCCATCGGCAAGCTGTGTGGGAGATGCAAGGCAACAGCGCTCTCTCGGACAGCACT 359
Qf 1092 TCAGCTTCAATATCATCTTTAGTGCAGAGTTCGAAAACGTTTCTGTGTAGATGGGCTAATG 1151
Df 360 TCAGCTTCAATATCATCTTTAGTGCAGAGTTCGAAAACGTTTCTGTGTAGATGGGCTAATG 419
Qf 1152 GAATGGCAAGTGTGAGCTTTCTCCCTCTTTCACCTCTTCCCTTCCAAATCTTCAATGACA 1211
Df 420 GAATGGCAAGTGTGAGCTTTCTCCCTCTTTCACCTCTTCCCTTCCAAATCTTCAATGACA 479
Qf 1212 GACAGTGTCTTGTGATATAAGCGCTGTGAAATAAGGTAATGCAA 1257
Df 480 GACAGTGTCTTGTGATATAAGCGCTGTGAAATAAGGTAATGCAA 524

```

RESULT 7
AI854700/c

FEATURES

LOCUS
 DEFINITION UI-M-BHO-akc-d-12-0-UI.s1 NIH BMAP_M_S1 Mus musculus cDNA clone
 ACCESSION UI-M-BHO-akc-d-12-0-UI 3', mRNA sequence.
 VERSION A1854700
 KEYWORDS A1854700.1 GI:5498606
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 499)
 Bonaudo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 normalized basal ganglia library cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
 available by the means that is soon to be determined. When NIH
 determines the means for distribution of the BMAP cDNA clones, this
 record will be updated accordingly when that means is determined.
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES
 source Location/Qualifiers
 1..499
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BHO-akc-d-12-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_M_S1"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP M_S1 library is a subtracted library derived from
 a mixture of normalized libraries from ten regions of the
 mouse brain (cerebellum, brain stems, olfactory bulbs,
 hypothalamus, cortex, amygdala, basal ganglia, pineal
 gland, striatum, hippocampus). The driver used for
 subtraction consisted of a pool of 20,000 cDNA clones
 obtained from non-normalized and normalized libraries of
 these ten regions of the mouse brain.
 TAG TISSUE=basal-ganglia
 TAG_LIB=NIH_BMAP_M_S1
 TAG_SEQ=TGATAC"

ORIGIN
 Query Match 37.1%; Score 473.8; DB 9; Length 499;
 Best Local Similarity 99.2%; Pred. No. 1.3e-61;
 Matches 497; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
 QY 772 GAGCAGCAAGAGCGCTTCCAGTTCTTAGAGCAGAGTACGCTACTATCAC 831
 Db 499 GAGCAGCAAGAGCGCTTCCAGTTCTTAGAGCAGAGTACGCTACTATCAC 440
 QY 832 TGCAGGACTGCAAAATCCGTTGGGAGAGCGCTATGTGTGTGTGTCGAGGCGCACCAGT 891
 Db 439 TGCAGGACTGCAAAATCCGTTGGGAGAGCGCTATGTGTGTGTGTCGAGGCGCACCAGT 380

QY 892 AAGGTGTTACTTCAAAACAGTTCTGCCAGTGTGTGAGAAAATCTTACAAACCTTACAGAGT 951
 Db 379 AAGGTG-TACTTCAAAACAGTTCTGCCAGTGTGTGAGAAAATCTTACAAACCTTACAGAGT 321
 QY 952 GGAGGCATCATCCTGTCAAAGTTGTAAAGAACTAGATGTGCCTGCCAGTCACATTTCG 1011
 Db 320 GGAGGCATCATCCTGTCAAAGTTGTAAAGTAACTAGATGTGCCTGCCAGTCACATTTCG 261
 QY 1012 CCACGTGGACCCCTAAACGCCCCCATCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACG 1071
 Db 260 CCACGTGGACCCCTAAACGCCCCCATCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACG 201
 QY 1072 CCGTCTCTGGCAGCAGCCTTTCAGCTTCAATACATCATTTAGTGAGAGTCGAAAACGTT 1131
 Db 200 CCGTCTCTGGCAGCAGCCTTTCAGCTTCAATACATCATTTAGTGAGAGTCGAAAACGTT 141
 QY 1132 TCTGCTAGATGGGGCTAATGGAATGGCAAGTGAAGCTTTCCTCCCTCTTCCCTCTTCCC 1191
 Db 140 TCTGCTAGATGGGGCTAATGGAATGGCAAGTGAAGCTTTCCTCCCTCTTCCCTCTTCCC 81
 QY 1192 TTTCCAAATTTCTTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAGGTA 1251
 Db 80 TTTCCAAATTTCTTCATGACAGACAGTGT-TACTTGGATATAAAGCCTGTGAATAAAGGTA 22
 QY 1252 TTGCAAAACAAAAA 1272
 Db 21 TTGCAAAACAAAAA 1

RESULT 8
 BB703259
 LOCUS BB703259
 DEFINITION BB703259 RIKEN full-length enriched, in vitro fertilized eggs Mus
 musculus cDNA clone 7420449J15 3', mRNA sequence.
 ACCESSION BB703259
 VERSION BB703259.1 GI:16052094
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 491)
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
 Tanaka,T., Tonaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES

Location/Qualifiers
 1. 491
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="7420449J15"
 /sex="female"
 /tissue_type="in vitro fertilized eggs"
 /dev_stage="egg"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTTAATAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

ORIGIN

Query Match 35.1%; Score 448.4; DB 10; Length 491;
 Best Local Similarity 98.2%; Pred. No. 8.8e-58;
 Matches 485; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
 QY 763 CAAGCAGGAGCAGGACAGGAGCGCTCGTTCCAGTTCTTAGACGAGAGTACGGC 822
 Db 1 CAAGCAGGAGCAGGACAGGAGCGCTCGTTCCAGTTCTTAGACGAGAGTACGGC 60
 QY 823 TACTATCACTCAAGGACTGCAAAATCCGGTGGGAGAGCGCTATGTGTGGTGTGCAG 882
 Db 61 TACTATCACTGCAA-GACTGCAAAATCCGGTGGGAGAGCGCTATGTGTGGTGTGCAG 119
 QY 883 GGCACCAAGTGGTTACTTCAACAGTTCTGCGAGTGTGTGAGAAATCTCAACCC 942
 Db 120 GGCACCAAGTGGT-TACTTAAACAGTTCTGCGAGTGTGTGAGAAATCTCAACCC 178
 QY 943 TTACAGAGTGAGGACATCACTGTCAAGTTGTAAAGCACTAGATGTGCTGCCAGT 1002
 Db 179 TTACAGAGTGAGGACATCACTGTCAAGTTGTAAAGCACTAGATGTGCTGCCAGT 238
 QY 1003 CAGATTTCCGACGTGGACCTTAAACGCCGCCCATCGGAAGACTTGTGTGGAGATGCAA 1062
 Db 239 CAGACTTCCGACGTGGACCTTAAACGCCGCCCATCGGAAGACTTGTGTGGAGATGCAA 298
 QY 1063 GGACAAACCCCTGTCTCGGACAGACCTTCAAGTTCAATATCATCATTTAGTAGAGTC 1122
 Db 299 GGACAAACCCCTGTCTCGGACAGACCTTCAAGTTCAATATCATCATTTAGTAGAGTC 358
 QY 1123 GAAACGTTTCTGTAGATGGGGCTAATCGAATGACAGAGTGTCTTCCCTCTTCA 1182
 Db 359 GAAACGTTTCTGTAGATGGGGCTAATCGAATGACAGAGTGTCTTCCCTCTTCA 418

QY 1183 CCTCTCCCTTCCAAATCTTCATGACAGACAGTGTACTTGGATATAAGCCTGTGAA 1242
 Db |||||

419 CCTCTCCCTTCCAAATCTTCATGACAGACAGTGTACTTGGATATAAGCCTGTGAA 477
 |||||

QY 1243 TAAAGGTATTGCA 1256
 |||||

Db 478 TAAAGGTATTGCA 491
 |||||

RESULT 9

BB704648

LOCUS

DEFINITION

BB704648 RIKEN full-length enriched, in vitro fertilized eggs Mus

musculus cDNA clone 7420466L07 3', mRNA sequence.

ACCESSION

BB704648

VERSION

BB704648.1 GI:16053483

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 491)

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,

Hayatsu, N., Hiramoto, K., Hiraoka, F., Hirozane, T., Imotani, K.,

Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,

Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,

Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Tanaka, T., Tomaru, A., Toya, I., Watahiki, A., Yasunishi, A.,

Muramatsu, M., and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.

2001)

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.
 e mouse tissues.

FEATURES

Location/Qualifiers

1. 491

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="7420466L07"

/sex="female"

/tissue_type="in vitro fertilized eggs"

/dev_stage="egg"


```

Best Local Similarity 86.1%; Pred. No. 5.3e-50;
Matches 464; Conservative 0; Mismatches 71; Indels 4; Gaps 3;

QY 713 TGGCTCAGGACCCCGGTGATTCGATGCGCTCCAGACCGGCTCCCGCAAGACGCG 772
Db 542 TGGNCTNAGACCCCGGTATTCGATGCGCTCCAGACCGGCTCCCGCAAGACGCG 483
QY 773 AGCAGGACAAAGGAGCGCTGGTTTCCAGATTCCTTAGAGCAGAACTACGG--CTACTATCA 830
Db 482 AGNACAGGAGCGCGCTGCNTTTTCCAGCTCTTAGAGCAGAACTAGCNCCTACTATCA 423
QY 831 CTGCAAGGACTGCAAAATCCCGTGGGAGAGCGCTATGTG-TGGTGTGTGAGGCGCACCA 889
Db 422 CTGCAAGGAGCTGCAAAATCCCGTGGGAGAGCNCCTATGTGTGTGTGAGGCGCACCC 363
QY 890 GTAAGGTGTACTTCAACAGTCTCGCGAGTGTGTGAGAACTCTCAACCTTACAGA 949
Db 362 AGTAAAGTGTACTTCAACAGTCTCGCGAGTGTGTGAGAACTCTCAACCTTACAGA 303
QY 950 GTGAGGACATCCTGTCAAAAGTTGTAAGAACTAGATGTGCTGCCAGTCAGATTT 1009
Db 302 GTGAGGACATCCTGTCAAAAGTTGTAAGAACTAGATGTGCTGCCAGTCAGATTT 243
QY 1010 CGCAGTGTGACCTTAAACGCCCCCATCGGCAAGACTGTGTGGAGATGCAAGGACAAA 1069
Db 242 CGCAGTGTGACCTTAAACGCCCCCATCGGCAAGACTGTGTGGAGATGCAAGGACAAA 183
QY 1070 CGCTGTGCTCGACAGCAGCTTCAAGTGTGAGAACTAGATGTGCTGCCAGTCAGATTT 1129
Db 182 CGCTGTGCTCGACAGCAGCTTCAAGTGTGAGAACTAGATGTGCTGCCAGTCAGATTT 123
QY 1130 TTTCTGTAGATGGGCTAATGAAATGGAATGCAAGTGTGCTTCTCCCTTCACTCTTC 1189
Db 122 TTTCTGTAGATGGGCTAATGAAATGGAATGCAAGTGTGCTTCTCCCTTCACTCTTC 63
QY 1190 CTTTCAAAATCTCATGACAGACAGTGTACTTGGATATAAAGCTGTGMAATAAAG 1248
Db 62 CTTTCAAAATCTCATGACAGACAGTGTACTTGGATATAAAGCTGTGMAATAAAG 5

RESULT 13
BE946858
LOCUS
DEFINITION
UT-M-BH3-awu-b-08-0-UI.s1 NIH BMAP M S4 Mus musculus cDNA clone
UT-M-BH3-awu-b-08-0-UI 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 450)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: m85@mail.nih.gov
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements

```

Seq primer: M13 Forward
POLYA=No.

FEATURES

source
1.450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-awu-b-08-0-UI"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP M S4"
/notes="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH BMAP M S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH BMAP M S4,
NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,
NIH BMAP M S2, NIH BMAP M S1. The subtracted library
(NIH BMAP M S4) was constructed as follows: PCR amplified
cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and
NIH BMAP M S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH BMAP M S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_SEQ=None found"

ORIGIN

Query Match 30.8%; Score 393; DB 10; Length 450;
Best Local Similarity 99.7%; Pred. No. 1.9e-49;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 GCGCGGAGCGCGGAGCAGCCATGTTCCCGGAGCAGCAGTTCACCCCTGCCGCA 62
Db 57 GCGCGGAGCGCGGAGCAGCCATGTTCCCGGAGCAGCAGTTCACCCCTGCCGCA 116
QY 63 TCCTTATCCGAGGCCACCAAGCCGGGATGCTGGAGTTCGGAGCCAGGGCTGCCG 122
Db 117 TCCTTATCCGAGGCCACCAAGCCGGGATGCTGGAGTTCGGAGCCAGGGCTGCCG 176
QY 123 ACCCGCGCCCTCTCTTCTCCCGGCTACACAGCTCATCGCGCGAGTACGTCGA 182
Db 177 ACCCGCGCCCTCTCTTCTCCCGGCTACACAGCTCATCGCGCGAGTACGTCGA 236
QY 183 CAGCCACAGCGGCGCAGCTCATGCTCATGCTGTGTGCGGATGGGTCCCGGTGCGT 242
Db 237 CAGCCACAGCGGCGCAGCTCATGCTCATGCTGTGTGCGGATGGGTCCCGGTGCGT 296
QY 243 CAGCCGTGAGCGTGGCGTGCAGTGAACCGCGCGGAGCGCTCGGTGAGTGTCACT 302
Db 297 CAGCCGTGAGCGTGGCGTGCAGTGAACCGCGCGGAGCGCTCGGTGAGTGTCACT 356
QY 303 CCGCGCGCCGCGCAGCTGCAGGCTGCCAGCCAGCCAGCCGCGGATCGGGTTC 362
Db 357 CCGCGCGCCGCGCAGCTGCAGGCTGCCAGCCAGCCAGCCGCGGATCGGGTTC 416
QY 363 CTGTCAACCCCGTGGCCAGCCCGCGCGCGGAGAGA 396
Db 417 CTGTCAACCCCGTGGCCAGCCCGCGCGCGGAGAGA 450

[illegible]

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>,
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

FEATURES	source
Location/Qualifiers	
1. .419	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6J"	
/db_xref="taxon:10090"	
/clone="7420464A04"	
/sex="female"	
/tissue_type="in vitro fertilized eggs"	
/dev_stage="egg"	
/lab_host="DH30B"	
/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"	
/note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken	

contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using thermostable thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATCTTCGAGTTTAAATTAATCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+), after bulk excision from Lambda
ETC I phagein sites. 5', end: BamHI, 3', end: BamHI".

[illegible]

RESULT 15	521 bp	linear	EST 11-OCT-2001
BB703869	BB703869	RIKEN full-length enriched, in vitro fertilized eggs Mus	
LOCUS	BB703869	musculus cDNA clone 7420457C21 3'	
DEFINITION	BB703869	musculus cDNA clone 7420457C21 3'	

BB703869
 BB703869.1 GI:16052704
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 521)
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,
 Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
 Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sugabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Tanaka, T., Tomaru, A., Toya, T., Wataniki, A., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.)
 TITLE

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 15:24:32 ; Search time 5112 Seconds
(without alignments)
10827.280 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagggggcgagcgcgga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1245.4	97.5	1260	10	AY191415	AY191415 Mus muscu
2	562.2	75.3	1280	10	AY283175	AY283175 Rattus no
3	799.4	62.6	3987	10	AY193889	AY193889 Mus muscu
4	799.4	62.6	147464	2	AC122733	AC122733 Mus muscu
5	799.4	62.6	200535	10	AL671880	AL671880 Mouse DNA
6	689.8	54.0	168073	2	AC107686	AC107686 Mus muscu
7	628	49.2	235390	2	AC126519	AC126519 Rattus no
8	628	49.2	237695	2	AC125993	AC125993 Rattus no
9	628	49.2	262139	2	AC127083	AC127083 Rattus no
10	360	28.2	212848	2	AC108848	AC108848 Mus muscu
11	352	27.6	1275	9	AY191416	AY191416 Homo sapi
12	299.8	23.5	212848	2	AC108848	AC108848 Mus muscu
13	205.8	16.1	1052	5	AY283176	AY283176 Xenopus l
14	194.6	15.3	963	5	AY283177	AY283177 Takifugu
15	194.6	15.2	1084	5	AY283178	AY283178 Danio rer
16	147.6	11.6	138890	9	AC007970	AC007970 Homo sapi
17	147.6	11.6	151961	2	AC037443	AC037443 Homo sapi
18	147.6	11.6	204268	2	AC092163	AC092163 Homo sapi
19	134.6	10.5	4138	9	AY193890	AY193890 Homo sapi
20	134.6	10.5	187578	9	AC096952	AC096952 Homo sapi
21	90.6	7.1	187278	2	BX571888	BX571888 Danio rer
22	90.6	7.1	249791	5	BX537133	BX537133 Zebrafish
23	73.2	5.7	694	9	HS4335296	AJ335296 Homo sapi
24	72.2	5.7	125020	9	AF429315	AF429315 Homo sapi
25	64.8	5.1	144000	9	AC123764	AC123764 Homo sapi
26	63.2	4.9	125020	9	AF429315	AF429315 Homo sapi
27	62.6	4.9	1393	11	PM11H12G	AL684264 Penicilli
28	62	4.9	136551	2	AC048354	AC048354 Homo sapi
29	58.6	4.6	201858	2	AC115880	AC115880 Mus muscu
30	58.4	4.6	248550	1	SC0939120	AL939120 Streptomy
31	58.2	4.6	94058	2	AC141024	AC141024 Rattus no
32	58	4.5	92310	2	AC108380	AC108380 Pan trogl
33	57.8	4.5	2000	6	AX655393	AX655393 Sequence
34	57.2	4.5	143406	10	AC092857	AC092857 Rattus no
35	57.2	4.5	230329	2	AC142070	AC142070 Rattus no
36	56.6	4.4	245210	2	AC137771	AC137771 Homo sapi
37	56.2	4.4	257784	2	AC118627	AC118627 Mus muscu
38	56	4.4	247641	5	BX470223	BX470223 Zebrafish
39	56	4.4	248640	2	BX569785	BX569785 Danio rer
40	56	4.4	259098	2	BX548044	BX548044 Danio rer
41	55.6	4.4	991	11	PM12H12B	AL684455 Penicilli
42	55.6	4.4	151349	10	AL355176	AL355176 Mouse DNA
43	55	4.3	745	9	HS4335067	AJ335067 Homo sapi
44	54.6	4.3	183305	2	BX640474	BX640474 Danio rer
45	54.4	4.3	976	11	PM12D6G	AL684370 Penicilli

ALIGNMENTS

RESULT 1
AY191415
LOCUS AY191415 1260 bp mRNA linear ROD 12-MAR-2003
DEFINITION Mus musculus zygote arrest 1 (Zar1) mRNA, complete cds.
ACCESSION AY191415
VERSION AY191415.1 GI:27808689
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1260)
AUTHORS Wu,X., Viveiros,M.M., Eppig,J.J., Bai,Y., Fitzpatrick,S.L. and Matzuk,M.M.
TITLE Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for

MEDLINE	22811438
PUBMED	12773403
REFERENCE	2 (bases 1 to 1280)
AUTHORS	Wu,X. and Matczuk,M.M.
TITLE	Direct Submission
JOURNAL	Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
FEATURES	Location/Qualifiers
source	1..1280
	/organism="Rattus norvegicus"
	/mol_type="mRNA"
	/db_xref="taxon:10116"
	/chromosome="14"
gene	1..1280
	/gene="Zar1"
CDS	28..1113
	/gene="Zar1"
	/note="maternal factor"
	/codon_start=1
	/product="zygote arrest 1"
	/protein_id="AAP37037.1"
	/db_xref="GI:30908933"
	/translation="MFPASTPHCPHPYPTAAKAGDGRFARGCEPEPSLPGLVR RLAASPDARPGSQRPSPARAGRPRSVRSRDAAVQNVPRDASVQSCLGRRILQPGR GRPAPTTRFEPEVAVMKAVPQSEGDVQAEGDQEPFPREDPDVSAAQMSEF GSSEPPAVBMADPDSVAADRASPQSTEQDKERLRFQLEKYGYHCKDCNIRW ESAVWCVCQTSKVYFKQFCRVCKEKSYPVRVEDITCQSKRTFCACPVRLRHVDPKR PHRQDLGCRKDXRLSCDSIFSFKYII"
ORIGIN	
Query Match	75.3%; Score 962.2; DB 10; Length 1280;
Best Local Similarity	87.7%; Pred. No. 1.9e-189;
Matches 1110; Conservative	0; Mismatches 143; Indels 12; Gaps 5;
QY	3 GGCGGGCGAGCGCGGACGCCATCTGTTC CGGCGAGACAGTTCACCCCTGCCGC GA 62
Dd	3 GGCGGGCGAGCGCGGACGCCATCTGTTC CGGCGAGACAGCGCCCATCCCGCGCA 62
QY	63 TCCTTTATCCG --- AGGCCAACCAAGCGGGGATGGCTGGAGGTTCCGAGCCAGGGGCTG 119
Dd	63 TCCTTACC CGCCACGCGAGCAAGCGGGATGGCTGGAGGTTTGAGCGCAGGGGCTG 122
QY	120 CCGACCGCGCCCGCTCTTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTAGCT 179
Dd	123 CAGGCCCGAGCCCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCCGCGGAGTACTT 182
QY	180 CGACAGCACAGCGGGCGACAGCTCATGGCCCTCTGTTCGGGATGGGTCCC CGTCCGT 239
Dd	183 TGACAGCTATCAGCGAGCGAGCTCATGGCCCTGTCTCGCGAATGGGTCCC CGCGCGT 242
QY	240 CAGCAGCCGTGACGCTCGGTGACAGTGAAACCGCGCGGAGCGCTCGGTGAGTGTTTC 299
Dd	243 CAGCAGCCCGACGCTCGGTGACAGTGAAACCGCGCGGAGTGCCCTCGGTGAGTGTTTC 302
QY	300 ACTCGGCGCGCACGCTGCAGGCTGCGAGGTTGCCAGCCAGCGCCCGACGCCGATCCGG 359
Dd	303 GCTCGGCGCGCACACTGCAGCGCTGACGCGCGGAGCCAGCCCGACGCCCGGCTCG 362
QY	360 TTCCTGTCAACCGGTGGCCACGCGCGCGCGGGAGATCCCGCGATCTCGGAGACCGT 419
Dd	363 TTCCTGCAACCCCGCAGCCCGCGAGGGCGGGAGACCCCGCGATCTCTGGCGCACCGT 422
QY	420 AGCCCGCTTCTCGTCGTGACCTTCTGTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479
Dd	423 CGCCCTCTGA CTGCGCCGTGACCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 482
QY	480 CAGGACAGACACCCACGAAGGAGGAGGCGAGCCCGGCATCTCTCGGGGACCCGGGAACCGGA 539
Dd	483 CAGGACAGCCCGACGAAGGAGGAGGAGGAGACCGGACCCACCGGGGACCCGGGAACCCGA 542
QY	540 GCCGAGAGAGGTGGCCGCGGAGGAAGCGGTCCCCCAGCCGCGGAGCGAGGAGGCGCATGT 599

JOURNAL	Submitted (09-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 3987)
AUTHORS	Wu, X. and Matzuk, M.M.
TITLE	Direct Submission
JOURNAL	Submitted (21-JAN-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REMARK	Sequence update by submitter
COMMENT	On Jan 21, 2003 this sequence version replaced gi:27808693.
FEATURES	Location/Qualifiers
source	1..3987
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/strain="129/SvEv"
	/db_xref="taxon:10090"
gene	1..3987
	/gene="Zarl1"
mRNA	join(1..801,2574..2666,2745..2819,3697..3987)
	/gene="Zarl1"
	/product="zygote arrest 1"
exon	1..801
	/gene="Zarl1"
	/number=1
	/product="zygote arrest 1"
	/protein_id="AA024708.1"
	/db_xref="GI:27808694"
CDS	join(28..801,2574..2666,2745..2819,3697..3987)
	/gene="Zarl1"
	/note="oocyte-specific"
	/codon_start=1
	/product="zygote arrest 1"
	/protein_id="AA024708.1"
	/db_xref="GI:27808694"
	/translation="MPPASTFHPHPYPPQATKAGDWMRFQARGCRPAPPFLPGYRQLMAEYVDSHQALHMLLSRMRPSVSRDAVONPRDASVCSLGRRTLOPAGCRASPDARSQOPRGHAGRSPRSQWTVAPFSVTFGLSSLSVACGRQTPTKGEGSPASSGTRPREYVARKAVQPRSEGDVQAAGAGHEQPPEDRNSVAAMQSEPGSEPCPAEMAQDFQSDAPDRQKSPQSTQBQKRLRFQLEKQIGYHCKDKIRWESAYVMQGTSTKVYKPCFRCVCKESYNPVEDITCQSKRTRCAGVRLRHVDPKRPHRQDLGCRKDKRLSCDSTFSFKVII"
exon	2574..2666
	/gene="Zarl1"
	/number=2
exon	2745..2819
	/gene="Zarl1"
	/number=3
exon	3697..3987
	/gene="Zarl1"
	/number=4
ORIGIN	
Query Match	62.6%; Score 799.4; DB 10; Length 3987;
Best Local Similarity	99.9%; Pred. No. 1.2e-155;
Matches	800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	3 GGCGGGGAGCGCGGAGCAGCCATGTTCCCGGGAGCAGCTTCCACCCCTGCGCGCA 62
Db	3 GGCGGGGAGCGCGGAGCGCACCCTATGTTCCCGGGAGCAGCTTCCACCCCTGCGCGCA 62
Qy	63 TCCTTATTCGAGGCGCCACCAAGACCGGGATGCTCGAGGTTTCGGAGCCAGGGGTGCGG 122
Db	63 TCCTTATTCGAGGCGCCACCAAGACCGGGATGCTCGAGGTTTCGGAGCCAGGGGTGCGG 122
Qy	123 ACCGGGCCCCCTCTCTCTCCCGGCTACAGACGCTCAGTGGCGCGGAGTACGTGCA 182
Db	123 ACCGGGCCCCCTCTCTCTCCCGGCTACAGACGCTCAGTGGCGCGGAGTACGTGCA 182
Qy	183 CAGCCACAGCGGCGACAGCTCATATGCGCCCTGTGTGTCGCGGATGGGTGCCCGGTTCAG 242
Db	183 CAGCCACAGCGGCGACAGCTCATATGCGCCCTGTGTGTCGCGGATGGGTGCCCGGTTCAG 242
Qy	243 CAGCGGTGAGCTCGGCTGACGTGACCGCGCGCGAGCGCTCGGTGACGTGTTCACT 302
Db	243 CAGCGGTGAGCGCTCGGCTGACGTGACCGCGCGCGAGCGCTCGGTGACGTGTTCACT 302
Qy	303 CGGCGCGGACCGCTGACGCTTCGAGGGTGGCGAGCGCAGCGCCCGGATCGGGTTC 362

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 147464)

TITLE JOURNAL

REFERENCE AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, B., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Granel, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (02-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 2, 2003 this sequence version replaced gi:21206317.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

TITLE JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L17506

Center clone name: 506_B_15

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 140214 bases at least Q40
Consensus quality: 143454 bases at least Q30
Consensus quality: 144691 bases at least Q20
Insert size: 144000; agarose-fp
Insert size: 145464; sum-of-contigs
Quality coverage: 4.6 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 653: contig of 653 bp in length
* 654 753: gap of 100 bp
* 754 1934: contig of 1181 bp in length
* 1935 2034: gap of 100 bp
* 2035 31996: contig of 29962 bp in length
* 31997 32096: gap of 100 bp
* 32097 32986: contig of 890 bp in length
* 32987 33086: gap of 100 bp
* 33087 33811: contig of 725 bp in length
* 33812 33911: gap of 100 bp
* 33912 35429: contig of 1518 bp in length
* 35430 35529: gap of 100 bp

* 35530 38157: contig of 2628 bp in length
* 38158 38257: gap of 100 bp
* 38258 41538: contig of 3281 bp in length
* 41539 41638: gap of 100 bp
* 41639 43241: contig of 1603 bp in length
* 43242 43341: gap of 100 bp
* 43342 46930: contig of 3589 bp in length
* 46931 47030: gap of 100 bp
* 47031 51340: contig of 4310 bp in length
* 51341 51440: gap of 100 bp
* 51441 56135: contig of 4695 bp in length
* 56136 56235: gap of 100 bp
* 56236 61691: contig of 5456 bp in length
* 61692 61791: gap of 100 bp
* 61792 70053: contig of 8262 bp in length
* 70054 70153: gap of 100 bp
* 70154 76663: contig of 6510 bp in length
* 76664 86117: contig of 9354 bp in length
* 86118 86217: gap of 100 bp
* 86219 94174: contig of 7957 bp in length
* 94175 94274: gap of 100 bp
* 94275 108523: contig of 14249 bp in length
* 108524 123796: contig of 15173 bp in length
* 123797 123896: gap of 100 bp
* 123897 146961: contig of 23065 bp in length
* 146962 147061: gap of 100 bp
* 147062 147464: contig of 403 bp in length.
FEATURES
Location/Qualifiers
source
1..147464
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-506B15"
/clone_lib="RPC1-24 Male Mouse BAC"
misc_feature
1..653
/note="assembly_fragment"
misc_feature
754..1934
/note="assembly_fragment"
misc_feature
2035..31996
/note="assembly_fragment"
misc_feature
32097..32986
/note="assembly_fragment"
misc_feature
33087..33811
/note="assembly_fragment"
misc_feature
33912..35429
/note="assembly_fragment"
misc_feature
35530..38157
/note="assembly_fragment"
misc_feature
38258..41538
/note="assembly_fragment"
misc_feature
41639..43241
/note="assembly_fragment"
misc_feature
43342..46930
/note="assembly_fragment"
misc_feature
47031..51340
/note="assembly_fragment"
misc_feature
51441..56135
/note="assembly_fragment"
misc_feature
56236..61691
/note="assembly_fragment"
misc_feature
61792..70053
/note="assembly_fragment"
misc_feature
70154..76663
/note="assembly_fragment"
misc_feature
76764..86117
/note="assembly_fragment"
misc_feature
86218..94174
/note="assembly_fragment"
misc_feature
94275..108523
/note="assembly_fragment"
misc_feature
108624..123796

```
/note="assembly_fragment"
123897..146961
/note="assembly_fragment"
147062..147464
/note="assembly_fragment
clone_end:T7
vector_side:right"

ORIGIN
Query Match 62.6%; Score 799.4; DB 2; Length 147464;
Best Local Similarity 99.9%; Pred. No. 1.1e-155;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCGGCGAGCGCGGCGAGCCACCATGTTCCCGGCGAGCAGTTCACCCCTGCCCGCA 62
Db 127364 GCGCGGCGAGCGCGGCGAGCCACCATGTTCCCGGCGAGCAGTTCACCCCTGCCCGCA 127423

QY 63 TCCTTATCCGAGCGCCACCAAGCCGGGATGGCTGGAGGTTTCGAGCCAGGGGTGCGG 122
Db 127424 TCCTTATCCGAGCGCCACCAAGCCGGGATGGCTGGAGGTTTCGAGCCAGGGGTGCGG 127483

QY 123 ACCCGGCCCCCTCTCTTCTCCCGGCTACAGAGCTATGCGCGGCGGAGTACGTGA 182
Db 127484 ACCCGGCCCCCTCTCTTCTCCCGGCTACAGAGCTATGCGCGGCGGAGTACGTGA 127543

QY 183 CAGCCACAGCGGCGACAGCTCATGSCCTGCTGTGCGGATGGTTCGCCGTCCGTGAG 242
Db 127544 CAGCCACAGCGGCGACAGCTCATGSCCTGCTGTGCGGATGGTTCGCCGTCCGTGAG 127603

QY 243 CAGCGGTGACGCTGCGGTGAGTGAACCCGCGCGAGCGCTCGGTGAGTGTCACT 302
Db 127604 CAGCGGTGACGCTGCGGTGAGTGAACCCGCGCGAGCGCTCGGTGAGTGTCACT 127663

QY 303 CGGCGCGCGCAGCTGCGAGCTGCGAGGTCGAGCCAGCCCGCGAGCCCGGATCGGTTTC 362
Db 127664 CGGCGCGCGCAGCTGCGAGCTGCGAGGTCGAGCCAGCCCGCGAGCCCGGATCGGTTTC 127723

QY 363 CTGTCAACCCCGTGGCGAGCGCGCGCGGAGATCCCGCGATCTTGGCAGACCGTAGC 422
Db 127724 CTGTCAACCCCGTGGCGAGCGCGCGGAGATCCCGCGATCTTGGCAGACCGTAGC 127783

QY 423 CCGCTTCTCGTCCGTGACCTTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
Db 127784 CCGCTTCTCGTCCGTGACCTTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 127843

QY 483 GAGACACCAACGAGGAGAGGAGCGCGGATCTCTGCGGACCGGAGACCGGAGCC 542
Db 127844 GAGACACCAACGAGGAGAGGAGCGCGGATCTCTGCGGACCGGAGACCGGAGCC 127903

QY 543 GAGAGAGTGGCGCGAGGAAAGCGGTCCCGCGCGAGCGGAGGCGGATGTTCA 602
Db 127904 GAGAGAGTGGCGCGAGGAAAGCGGTCCCGCGCGAGCGGAGGCGGATGTTCA 127963

QY 603 GGTGTCAGGCGAGCGCGGTGGAGCAGCAGCCACCGGAGGACCGGAAAGTGTGCG 662
Db 127964 GGTGTCAGGCGAGCGCGGTGGAGCAGCAGCCACCGGAGGACCGGAAAGTGTGCG 128023

QY 663 GCGGATGCACTGAGCTGGAGCGGAGGAGGATGTCCTGCGCGAGAGTGCCTCAGGA 722
Db 128024 GCGGATGCACTGAGCTGGAGCGGAGGAGGATGTCCTGCGCGAGAGTGCCTCAGGA 128083

QY 723 CCGCGGTGATTCGAGTCCCTCTGAGACCAAGGCGCTCCCGCAAGGACCGGAGGAGCA 782
Db 128084 CCGCGGTGATTCGAGTCCCTCTGAGACCAAGGCGCTCCCGCAAGGACCGGAGGAGCA 128143

QY 783 GGAGCGCCTCGGTTTCAGTT 803
Db 128144 GGAGCGCCTCGGTTTCAGTT 128164

RESULT 5
AL671880 200535 bp DNA linear ROD 27-NOV-2002
LOCUS
```



```

Qy 660 GCGGCGGATGAGTCTGAGCGCTGGAGGAGGAGCCATGTCTGTCGCGAGAGATGGCTCA 719
    |||
Db 2681.2 GCGGCGGATGAGTCTGAGCGCGGAGTGAGGAGCCACTCTGCTGTCGAGATGGCTCA 26753
    |||

Qy 720 GGACCCCGGTGATTCGGATGCCCTCGAGACGAGCCCTCCCGGCAAGACGAGGACGAGA 779
    |||
Db 26752 GGACCCCGGTGATTCGGATGCCCTCGAGACGAGCCCTCCCGGCAAGACGAGGACGAGA 26693
    |||

Qy 780 CAAGGAGCGCTGGCTTTCCAGTT 803
    |||
Db 26692 CAAGGAGCGCTGGCTTTCCAGTT 26669
    |||

RESULT 8
AC125993/c
LOCUS      Rattus norvegicus clone CH230-74L11, *** SEQUENCING IN PROGRESS
DEFINITION      ***
ACCESSION      AC125993
VERSION        AC125993.3 GI:30522839
KEYWORDS       HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
  1 (bases 1 to 237695)
    Muzny,D,Marie., Metzker,M, Lee., Abranzon,S., Adams,C., Alder,J.,
    Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
    Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
    Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
    Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
    Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
    Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
    Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
    Cleaveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
    Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
    Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
    Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
    Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
    Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
    Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
    Gebregiorge,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
    Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
    Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
    Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
    Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
    Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
    Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
    Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
    Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
    Lorensuheva,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
    Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
    Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
    Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
    Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
    Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
    Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
    Nwackeleme,O., Okwuonu,G., Olarpunsgoon,A., Pal,S., Parks,K.,
    Pasternak,S., Paul,H., Perez,A., Perez,L., Pfanckoch,C.,
    Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
    Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
    Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
    Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
    Sanders,W., Savery,G., Scherer,S., Scott,G., Shataman,S., Shen,H.,
    Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
    Sneed,A., Sodergren,S., Song,X.-Z., Sorelle,R., Sosa,J.,
    Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
    Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
    Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
    Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
    Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
    Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

```

```

Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dumm,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 237695)
Worley,K.C.
Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237695)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGIZ
Center clone name: CH230-74L11
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 229641 bases at least Q40
Consensus quality: 231444 bases at least Q30
Consensus quality: 232679 bases at least Q20
Estimated insert size: 242796; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 237695: contig of 237695 bp in length.
FEATURES
    source
        1..237695
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-74L11"
    misc_feature
        1..1905
        /note="wgs end extension
        clone end:T7"
    misc_feature
        complement(5394..6172)
        /note="clone boundary
        clone end:T7
        site:ECORI
        end sequence:BH339391"
    misc_feature
        complement(234275..235086)

```


Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 212848)

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 28, 2002 this sequence version replaced gi:18450109.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L21014

Center clone name: 300_K_5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 200069 bases at least Q40

Consensus quality: 206302 bases at least Q30

Consensus quality: 208218 bases at least Q20

Insert size: 209348; sum-of-coverage

Quality coverage: 5.7 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1

* 641: contig of 841 bp in length

* 741: gap of 100 bp

* 1427: contig of 686 bp in length

* 1527: gap of 100 bp

* 1528: contig of 1643 bp in length

* 3170: contig of 100 bp

* 3171: gap of 100 bp

* 3270: contig of 1126 bp in length

* 4396: gap of 100 bp

* 4496: gap of 100 bp

* 5909: contig of 1413 bp in length

* 6010: gap of 100 bp

* 7583: contig of 1574 bp in length

* 7683: gap of 100 bp

* 8975: contig of 1292 bp in length

* 9075: gap of 100 bp

* 10257: contig of 1182 bp in length

* 10357: gap of 100 bp

* 11608: contig of 1251 bp in length

* 11708: gap of 100 bp

* 13289: contig of 1581 bp in length

* 13389: gap of 100 bp

* 15335: contig of 1946 bp in length

* 15435: gap of 100 bp

* 16250: contig of 815 bp in length

* 16350: gap of 100 bp

* 17689: contig of 1319 bp in length

* 17769: gap of 100 bp

* 19907: contig of 2138 bp in length

* 20007: gap of 100 bp

* 21470: contig of 1463 bp in length

* 21570: gap of 100 bp

* 23086: contig of 1516 bp in length

* 23186: gap of 100 bp

* 24723: contig of 1537 bp in length

* 24823: gap of 100 bp

* 27746: contig of 2923 bp in length

* 27846: gap of 100 bp

* 29873: contig of 2027 bp in length

* 29973: gap of 100 bp

* 33212: contig of 3239 bp in length

* 33312: gap of 100 bp

* 61732: contig of 28420 bp in length

* 61832: gap of 100 bp

* 63732: contig of 1960 bp in length

* 63832: gap of 100 bp

* 67231: contig of 3339 bp in length

* 67331: gap of 100 bp

* 70922: contig of 3591 bp in length

* 71022: gap of 100 bp

* 76930: contig of 5908 bp in length

* 77030: gap of 100 bp

* 82623: contig of 5593 bp in length

* 82723: gap of 100 bp

* 87875: contig of 5152 bp in length

* 87975: gap of 100 bp

* 87976: contig of 5799 bp in length

* 93874: gap of 100 bp

* 93875: contig of 10331 bp in length

* 104205: gap of 100 bp

* 104306: contig of 14799 bp in length

* 119104: gap of 100 bp

* 119204: contig of 13095 bp in length

* 132299: gap of 100 bp

* 132300: contig of 16571 bp in length

* 148970: gap of 100 bp

* 149071: contig of 16644 bp in length

* 165715: gap of 100 bp

* 165815: contig of 17555 bp in length

* 183370: gap of 100 bp

* 183470: contig of 26822 bp in length

* 210291: gap of 100 bp

* 210292: contig of 2457 bp in length.

* 210392: Location/Qualifiers

FEATURES

source	1. .212848 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /clone="RP23-300K5" /clone_lib="RPCI-23 Female Mouse BAC" misc_feature 1. .641 /note="assembly_fragment" 742. .1427 /note="assembly_fragment" 1528. .3170 /note="assembly_fragment" 3271. .4396 /note="assembly_fragment" 4497. .5909 /note="assembly_fragment" 6010. .7583 /note="assembly_fragment" 7684. .8975 /note="assembly_fragment" 9076. .10257 /note="assembly_fragment" 10358. .11608 /note="assembly_fragment" 11709. .13289 /note="assembly_fragment" 13390. .15335 /note="assembly_fragment" 15436. .16250 /note="assembly_fragment" 16351. .17669 misc_feature	ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL FEATURES source gene CDS
--------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ACCESSION	AY191416
VERSION	AY191416.1 GI:27808691
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1275) Wu,X., Viveiros,M.M., Eppig,J.J., Bai,Y., Fitzpatrick,S.L. and Matzuk,M.M.
AUTHORS	Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition
TITLE	Nat. Genet. 33 (2), 187-191 (2003)
JOURNAL	22447938
MEDLINE	PUBMED
REFERENCE	2 (bases 1 to 1275) Wu,X., Bai,Y. and Matzuk,M.M. Direct Submission
AUTHORS	Submitted (04-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
TITLE	Location/Qualifiers
JOURNAL	1..1275
FEATURES	/organism="Homo sapiens" /mol_type="mrna" /db_xref="taxon:9606" /chromosome="4" /map="4p12" 1..1275 /gene="ZAR1" 1..1275 /gene="ZAR1" /codon_start=1 /product="zygote arrest 1" /protein_id="AA024707.1" /db_xref="GI:27808692" /translations="MAALGEVLDGVFPACPSCSYRYPYPAATKGKGAAGSGWQORGLCPASSPCSAAGSIPEPGCRLTAAVEFDSYORERLMAILAQVGPGLGPPARRAGSCDAVQSPRIDAAVQCISLGRTLORRADPESAPGAEGTTGGSFQSQPSSRGLEQSQVQNGAPRRFTVAIVYSLALRLTALEFGPAGQRGSADSGERGPPA RLQPSBEQVTWKAPRPPOSDDGAQAARASWSBPADGPDLPPREAQGEAARPS ALRSPGPGPAGRDGDREAAVAGEGPSRSPELKERLPFLFEQKYGYHKDCNTRWEAYVMVCQTNKVYFKQFCRTCKSKSYNPVEDITCSQCKQTRCCPKVLRLH VDKPRHRQDLCRCCKGRKLSCDSTFSFKVII"
source	
genbank	
cds	
origin	
Query Match	27.6%; Score 352; DB 9; Length 1275;
Best Local Similarity	64.0%; Pred.No.le-62;
Matches	708; Conservative 0; Mismatches 280; Indels 118; Gaps 7;
Qy	126 CGCGCCCCCTCTTCCTCCCGGTACAGACAGCTCATGGCGCGGAGTAGCTGCACAG 185
Dd	171 CGCGGCTCTGTTCTTCCTCCCGGCTCGGGCGGCTACGCCGCCCGAGTACTTCGCACAG 230
Qy	186 CCACAGCGGGCACAGCTCATGGCCCTGCTGTTCGCGGATGGGTC----- 229
Dd	231 CTACCAGGGAGCGGCTCATGGCTCTCTTGCGCGAGTGGGGCGGGTCTCGGGCGCGCG 290
Qy	230 --CCGCTGCGTCAAGACCGCTGACGCTGCGGTGTCAGGTGAACCCGCGCGCGACGCTC 287
Dd	291 CGCGCGAGGGCCGCGAGCTCGACGCTGGGGGTTCAGGTGAGCGCGCGCGCATCGACGCGC 350
Qy	288 GGTGAGTGTTCATTCGGGCGCCGACAGCTGACGCTGACGCTGCGAGGTGCGAGCAGCCCCGA 347
Dd	351 GTTACAGTGTCTGCTGGGAGGCGACGCTGACGCGCGGCGCGCGACCCCGAGTCCCC 410
Qy	348 GCCTCGATCGGCTTCC----- 363
Dd	411 GCCCGCCC CGGGCCGAGGSCACACGGGTGGCGGTCTTCTTCACGAGCAGCATCCCG 470
Qy	364 -----TGTCAA CCGCTGTCACCGCGCGCGGAGATCCCGCGATCTCTGGCA 413
Dd	471 TCGAGGCTTGAGCAGGSCAGCCCGCAGAACCGGCGCGCGCGCCATGGGTTTCCCGCG 530

QY 414 GACCTAGCCCCGTTCTCGTCCGTGACCTCTGTGGCCCTCTCCCTCACTGAGGTTC 473
 Db 531 CACCGTCGCGGTGTTACTCGCCCTCGGCTTCGCGCGTCTACCGCTTCTTGAGGGGCC 590
 QY 474 GGGAGGC-----AGCGACACACCCACGAGGAGAGGGAGCCCGGCATC 518
 Db 591 CGGCCCGCGCGCGCGAGCAGAGGTTCGGGGCGTTCGGACGAGAGAGGGCGCGCGCC 650
 QY 519 CTCGGGACCCCGGAAACCGAGCGAGAGAGGTGCGCGGAGGAAAGGGTCCCCAGCC 578
 Db 651 CGCGCGCTTCAAGCCACAGAGAGGGGAGGTGTGACGAGAAAGAGCGCCCGCGCGCC 710
 QY 579 GC---GAAGCAGGAGGGCGATGTTTCAGGCTCGAGGCGAGCCCGGTGGAGC----- 628
 Db 711 GCAGTCGACGACGACGCGGAGGCCAGGCGCGAGTCGAGCGAGCTGGAGCAGCCGCG 770
 QY 629 -----AGCAGCCACCGAGGACCGGAGACCGAAGTGTGCGGGGATGCACTGGA 677
 Db 771 CGACGCTCCGAGCTGCGCGCGGAGAGGCGCCAGAGGCGGAGCGGTCCGCGGTCCGC 830
 QY 678 GCCTGGAGCGAGGAGCATGTCCTGCGCGAGAGATGCTCAGGACCCCGGTGATT---- 733
 Db 831 GCTAAGAGCGCGGGCAACTCGTTCGCGGGAGGGCGCGAGCGCGCGGACGAGC 890
 QY 734 -----CGATGCCCCCTCGACACCAAGGCTCCCCGAAAGCACGAGCAGGAGGAGCG 788
 Db 891 GGAGGCGCGCTCGCGGAGAGGGGCGCTCGCCACGAGCGCGGAGCTGGGCAAGGAGCG 950
 QY 789 CTGCGTTTCAGTCTTAGACGAGAAGTAGGCTACTATCAGTCGAGGAGCTGCAAAAT 848
 Db 951 GCTGCGCTTCCAGTCTTAGACGAGAATATGGCTATTACCACTGCAAGGACTGCAACAT 1010
 QY 849 CCGGTGGAGAGCGCCTATGTTGTTGTGTGAGGCGCAGGACCAAGTAAAGTGTACTTCAAC 908
 Db 1011 CCGTGGAGAGGCTTATGTTGTTGTGTATACAGGAACTAACAGGT-TTACTTTCAAC 1069
 QY 909 AGTTCTCCGAGTGTGTGAGAAATCTTACAACTTACAGAGTGGAGGACATCACCTGTC 968
 Db 1070 AGTTTTCGAACTTGTGAGAGTCTTATAACCTTACCGAGTGGAGGATATCACCTGTC 1129
 QY 969 AAGTTGTAAAGAACTAGATGTCCTGCGGAGTCAATTCGCGAGTGGAGCCCTAAAC 1028
 Db 1130 AAGTTGTAAACAAACGAGATGTTCTGCGCCAGTAAACTTGGCCAGCTGGAGCCCTAAAC 1189
 QY 1029 GCCCCATCGGAGACTTCTGTTGGAGATGCGAGGACAAACCGCTCTCTGGAGGCA 1088
 Db 1190 GGCCCAACCTCAGATGTTGTTGTTGATGCAAGGCAACCGCTCTCTGTGACGCA 1249
 QY 1089 CTTGAGTTCAATACATCAATTAG 1114
 Db 1250 CTTTACGTTCAATACATCAATTAG 1275

RESULT 12
 AC108848/c
 LOCUS Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered pieces.
 DEFINITION
 AC108848
 VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczek, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Meneus, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 212848)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczek, J., Levine, R., Lindblad-Toh, K., Liu, G., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 28, 2002 this sequence version replaced gi:18450109.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu

Center project name: L21014
 Center clone name: 300 K 5
 Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 20069 bases at least Q40
 Consensus quality: 206302 bases at least Q30
 Consensus quality: 208218 bases at least Q20
 Insert size: 209348; sum-of-contigs
 Quality coverage: 5.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

```

1 641: contig of 641 bp in length
* 642 741: gap of 100 bp
* 742 1427: contig of 686 bp in length
* 1428 1527: gap of 100 bp
* 1528 3170: contig of 1643 bp in length
* 3171 3270: gap of 100 bp
* 3271 4396: contig of 1126 bp in length
* 4397 4496: gap of 100 bp
* 4497 5909: contig of 1413 bp in length
* 5910 6009: gap of 100 bp
* 6010 7583: contig of 1574 bp in length
* 7584 7683: gap of 100 bp
* 7684 8975: contig of 1292 bp in length
* 8976 9075: gap of 100 bp
* 9076 10257: contig of 1182 bp in length
* 10258 10357: gap of 100 bp
* 10358 11608: contig of 1251 bp in length
* 11609 11708: gap of 100 bp
* 11709 13289: contig of 1581 bp in length
* 13290 13389: gap of 100 bp
* 13390 15335: contig of 1946 bp in length
* 15336 15435: gap of 100 bp
* 15436 16250: contig of 815 bp in length
* 16251 16350: gap of 100 bp
* 16351 17669: contig of 1319 bp in length
* 17670 17769: gap of 100 bp
* 17770 19907: contig of 2138 bp in length
* 19908 20007: gap of 100 bp
* 20008 21470: contig of 1463 bp in length
* 21471 21570: gap of 100 bp
* 21571 23086: contig of 1516 bp in length
* 23087 23186: gap of 100 bp
* 23187 24723: contig of 1537 bp in length
* 24724 24823: gap of 100 bp
* 24824 27746: contig of 2923 bp in length
* 27747 27846: gap of 100 bp
* 27847 29873: contig of 2027 bp in length
* 29874 29973: gap of 100 bp
* 29974 33212: contig of 3239 bp in length
* 33213 33312: gap of 100 bp
* 33313 61732: contig of 28420 bp in length
* 61733 61832: gap of 100 bp
* 61833 63792: contig of 1960 bp in length
* 63793 63892: gap of 100 bp
* 63893 67231: contig of 3339 bp in length
* 67232 67331: gap of 100 bp
* 67332 70222: contig of 3591 bp in length
* 70223 71022: gap of 100 bp
* 71023 76930: contig of 5908 bp in length
* 76931 77030: gap of 100 bp
* 77031 82623: contig of 5593 bp in length
* 82624 82723: gap of 100 bp
* 82724 87875: contig of 5152 bp in length
* 87876 87975: gap of 100 bp
* 87976 93774: contig of 5799 bp in length
* 93775 93874: gap of 100 bp
* 93875 104205: contig of 10331 bp in length
* 104206 104305: gap of 100 bp
* 104306 119104: contig of 14799 bp in length
* 119105 119204: gap of 100 bp
* 119205 132299: contig of 13095 bp in length
* 132300 132399: gap of 100 bp
* 132400 148970: contig of 16571 bp in length
* 148971 149070: gap of 100 bp
* 149071 165714: contig of 16644 bp in length
* 165715 165814: gap of 100 bp
* 165815 183369: contig of 17555 bp in length

```

```

* 183370 183469: gap of 100 bp
* 183470 210291: contig of 26822 bp in length
* 210292 210391: gap of 100 bp
* 210392 212848: contig of 2457 bp in length.

```

FEATURES

```

source
1. 212848
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /clone_lib="RPCI-23 Female Mouse BAC"
   /note="assembly_fragment"
1. 641
   /note="assembly_fragment"
742. 1427
   /note="assembly_fragment"
1528. 3170
   /note="assembly_fragment"
3271. 4396
   /note="assembly_fragment"
4497. 5909
   /note="assembly_fragment"
6010. 7583
   /note="assembly_fragment"
7684. 8975
   /note="assembly_fragment"
9076. 10257
   /note="assembly_fragment"
10358. 11608
   /note="assembly_fragment"
11709. 13289
   /note="assembly_fragment"
13390. 15335
   /note="assembly_fragment"
15436. 16250
   /note="assembly_fragment"
16351. 17669
   /note="assembly_fragment"

```

Query Match 23.5%; Score 299.8; DB 2; Length 212848;
 Best Local Similarity 72.6%; Pred. No. 6.8e-52;
 Matches 310; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

QY 3 GCGGGGAGGCGGCGGACCGACCCATGTTCCCGGCGAGCAGGTTCACCCCTGCGCGCA 62
Db 88270 GCGGGGAGGCGGCGGAAACGACCCATGTTCCCGGCGAGCAGGTTCACCCCTGCGCGCA 88211
QY 63 TCCTTATCCGAGCGCCACCAAGCGCGGATGGCTGGAGGTTCCGAGCAGGGGCTGCCG 122
Db 88210 TCCTTATCCGAGCGCCACCAAGCGCGGATGGCTGGAGGTTCCGAGCAGGGGCTGCCG 88151
QY 123 ACCGCGCCCCCTCTCTTCTCCCGGCTACAGACAGCTCATGGCCGGGAGTACGTGA 182
Db 88150 ACCGCGCCCCCTCTCTTCTCCCGGCTACAGACAGCTCATGGCCGGGAGTACGTGA 88091
QY 183 CAGCCACGCGGCGACAGCTCATGGCCCTGCTGCGGATGGTCCCGGTCGGTCCAG 242
Db 88090 CCGCCACGCGGCGACAGCTCATGGCCCTGCTGCGGATGGTCCCGGTCGGTCCAG 88031
QY 243 CAGCGGTGACGTGCGGTGCGAGTGAACCCCGCGCGAGCAACCGCTCGGTGAGTTCACT 302
Db 88030 CAGCGGTGACGTGCGGTGCGAGTGAACCCCGCGCGAGCAACCGCTCGGTGAGTTCACT 87971
QY 303 CGGCGCGCGACGCTGCGAGCCCTGCGAGGTGCGAGCCCGCGAGCCCGGATCGGGTTC 362
Db 87970 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87911
QY 363 CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGAGTCCCTCGCAGACCGTAGC 422
Db 87910 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87851
QY 423 CCGGTTTC 429
Db 87850 ACAGTTTC 87844

```

RESULT 13
AY2831176 1052 bp mRNA linear VRT 21-AUG-2003
LOCUS
DEFINITION Xenopus laevis zygote arrest 1 (Zarl) mRNA, complete cds.
ACCESSION AY2831176
VERSION AY2831176.1 GI:30908934
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 1052)
Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
Zygote arrest 1 (Zarl) is an evolutionarily conserved gene
expressed in vertebrate ovaries
Biol. Reprod. 69 (3), 861-867 (2003)
JOURNAL
MEDLINE
PUBMED
22811438
REFERENCE
2 (bases 1 to 1052)
Wu, X. and Matzuk, M.M.
Direct Submission
TITLE
Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
JOURNAL
FEATURES
source
1. .1052
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
gene
1. .1052
/gene="Zarl"
CDS
24. .911
/gene="Zarl"
/note="maternal factor"
/codon_start=1
/product="zygote arrest 1"
/protein_id="AAP37038.1"
/db_xref="GI:30908935"
translation="MYPAYNPYSYVILPNKMSRQKNYLSYGDYDNYORA
OLKALISQVNPNTPLRCRANTRDVGQVNPQDASVQCSGLPRTLLRRPGLRKPP
PEGSFASPTKTRFTTAIVSVAAGRLAPQDEGVNLEEKGEAVRSEGSGRQE
QKGDGEIKTEOMKMDTDEEAAAPQTRPKFQLEQYGYTHCKDNIRWESAYVWC
QETNIVYFKQFCRTCKSYNPYRVEDIMQSCQKTRCACPVKLRHVDPRPHRDLGG
RCKKRLSCDSTFSFKYII"
ORIGIN
Query Match 16.1%; Score 205.8; DB 5; Length 1052;
Best Local Similarity 80.1%; Pred. No. 2.4e-32;
Matches 254; Conservative 0; Mismatches 62; Indels 1; Gaps 1;
QY 796 TTCCAGTCTTACGACAGAGTACGCTACTATCAGTGAAGACTGCAAAATCCCGTGG 855
Db 594 TTCCAGTCTTACGACAGAGTACGCTACTATCAGTGAAGACTGCAAAATCCCGTGG 653
QY 856 GAGAGCGCCTATGTGTGGTGTGTGACAGGACCAAGTAAAGTGTACTTCAACAGTTCTG 915
Db 654 GAGAGCGCTACGTGTGTGTGTGACAGGACCAAGTAAAGTGTACTTCAACAGTTCTG 712
QY 916 CCGAGTGTGTGAATCTCAACCCCTTACAGAGTGGAGACATCACCTGTCAAAAGTTG 975
Db 713 CAGGACATGTCAAAATCTTAAATCCCTACCGTGTGGAAGACATCATGTGTCAAGCTG 772
QY 976 TAAAGAACTAGATGTGCGCCAGTCAAGTATTCGCCAGTGGACCTTAAAGCCCCCA 1035
Db 773 CAGGACATGTCAAAATCTTAAATCCCTACCGTGTGGAAGACATCATGTGTCAAGCTG 832
QY 1036 TCGGCAAGACTTGTGTGGAGATGCAAGGACAAACGCTGTCTCTGCGACAGACCTTCAG 1095
Db 833 CCGCCAGGATCTGTGTGGAGATGCAAGGACAAACGCTGTCTCTGCGACAGACCTTCAG 892
QY 1096 CTTCAAAATCATATTT 1112
|||||

Db 893 CTTCAAGTATATCATTT 909
RESULT 14
AY2831177 963 bp mRNA linear VRT 21-AUG-2003
LOCUS
DEFINITION Takifugu rubripes zygote arrest 1 (Zarl) mRNA, complete cds.
ACCESSION AY2831177
VERSION AY2831177.1 GI:30908936
KEYWORDS
SOURCE
ORGANISM
Takifugu rubripes (Fugu rubripes)
Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE
1 (bases 1 to 963)
Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
Zygote arrest 1 (Zarl) is an evolutionarily conserved gene
expressed in vertebrate ovaries
Biol. Reprod. 69 (3), 861-867 (2003)
JOURNAL
MEDLINE
PUBMED
22811438
REFERENCE
2 (bases 1 to 963)
Wu, X. and Matzuk, M.M.
Direct Submission
TITLE
Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
JOURNAL
FEATURES
source
1. .963
/organism="Takifugu rubripes"
/mol_type="mRNA"
/db_xref="taxon:31033"
gene
1. .963
/gene="Zarl"
CDS
1. .963
/gene="Zarl"
/note="maternal factor"
/codon_start=1
/product="zygote arrest 1"
/protein_id="AAP37039.1"
/db_xref="GI:30908937"
translation="MATYCDPEVDVSFYSSNYPMPYRPHRDAGHKYKSYLSHYGDT
SFAPNQRAQLKSILSQINPKLTPLRKANTKDVAQVNPQRDASVQCSIGPRTLLV
VRELRRRLKLPPEGTQKTEGEVRYPRTLAVYSPIAFRSVFLVETGDKRPAE
AQAEELPGQPKGNGENAGETNANLPQRKQSEDAQTADAGSGKARVRQF
LEQKYGYHRCNLEWESAYVWCQGTNKVYFKQFCRCKQKDFNRYVEDITCHVCN
KARCAEAETQRHVDPRPHRQDLGRCCKGKRLSCDSTFSFKYIV"
ORIGIN
Query Match 15.3%; Score 194.8; DB 5; Length 963;
Best Local Similarity 75.4%; Pred. No. 4.7e-30;
Matches 255; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
QY 776 AGGACAAGCAGCGCTGCGTTTCCAGTCTTACGACAGAGTACGCTACTACTCACTGCA 835
Db 626 AGGCAAAACGGGTGTCCTCCAGTCTTCCAGTCTTCCGAAACAGAAAGTACGCTACTCACTGCA 685
QY 836 AGGACTGCAAAATCCCGTGGAGAGCGCTATGTGTGTGTGTGTCAGGGCACCAGTAAGG 895
Db 686 GAGAATGCAACCTGGATGGAGAGCGCGTACGTTGTGTGTGCGTTTCAAGGCACCTAACAGG 745
QY 896 TGTACTTCAACAGTTCGCCGAGTGTGTGAGAAATCTCAACCCCTTACAGAGTGGAG 955
Db 746 T-TTACTTCAAGCAGTCTCTGTAGGAAATGCCAAAAGACTTTTAAACCCGTACCGCGTAGAG 804
QY 956 GACATCACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGCCCTGCGCTAGATTCGCCAC 1015
Db 805 GACATCAATGTACATGATGCAACAGGCCCGCTGTGCTGCGCGAGAAACGAGCGCCAC 864
QY 1016 GTGGACCCCTAAACGCCGCCCATCGGAAGACTTGTGTGGAGATGCAAGACAAACGCTGTG 1075
Db 865 GTTGACCCCAAGAGGCCCCACAGGACGCTGTGCGGAGGTGCNAGGCGGCGGTG 924
|||||

ORIGIN	Query Match	15.2%	Score 194.6;	DB 5;	Length 1084;
	Best Local Similarity	75.1%;	Pred. No. 5.2e-30;		
	Matches 256;	Conservative 0;	Mismatches 84;	Indels 1;	Gaps 1;
QY	774	GCAGGCAACAAGGAGCGCTCGCTTTCCAGTTCTTTAGACGAGAAGTACCGGTACTATCACTG	833		
DB	692	GAAGTCCAAGGCTCGTGTGAGATTTCAGTCTTTGGAGCAGAAGTATGGATTCTATCATTTG	751		
QY	834	CAAGGACTGC AAAATCCGGTGGGAGAGGCGCTATGTGTGTGTGTGTGTCAGGGCACCAGTAA	893		
DB	752	CAAAAGACTGC AACCTACGGTGGGAAAGTGCCTTATGTGTGTGTGTGTCCAGAGCAACAA	811		
QY	894	GGTGTACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTACAAACCCCTTTACAGAGTGG	953		
DB	812	GGT-TTATTTCAAGCAGTTCTGCAGAACATGCCAGAAATCATTTCAACCCATACCGGGTTG	870		
QY	954	AGGACATCACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGCCAGTCCAGATTTTCGCC	1013		
DB	871	AGGACATAGCATGTTCAGACTTTGCAGAAAGCTCCGCTGCACATGTTCTGTCAAGTCGCGTC	930		
QY	1014	ACGTGGACCTTAAACGCGCCCATCGGCAAGACTTTGTGTGGAGATGCAGAGCAAAACGCC	1073		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 13:40:31 ; Search time 577 Seconds
(without alignments)
9401.991 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence:

1 aagcgcgagcgagcgcgga.....acaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373663 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1265	99.1	1276	3 AAD00294	Aad00294 Mouse ooc
2	1234.2	96.6	1258	7 ABZ24589	Abz24589 Mouse ova
3	799.4	62.6	6873	7 ABZ24590	Abz24590 Mouse ova
4	753.8	59.0	4090	7 ABZ24591	Abz24591 Mouse ova
5	134.6	10.5	2075	7 ABZ24592	Abz24592 Human ova
6	58.8	4.6	114955	2 AAX53491	Aax53491 Human ade
7	57.8	4.5	2000	7 ADA71938	Ada71938 Rice gene
8	54.2	4.2	1032	7 ADA69587	Ada69587 Rice gene
9	52	4.1	114955	2 AAX53491	Aax53491 Human ade
10	51.6	4.0	987	7 ABZ37562	Abz37562 Streptomy
11	51.6	4.0	59816	7 ABZ37516	Abz37516 Streptomy
12	51.6	4.0	59816	7 ABZ37515	Abz37515 Streptomy
13	49.6	3.9	400	7 AAV44436	Aav44436 Mycobacte
14	49.6	3.9	400	2 AAV64545	Aav64545 M. tuberc
15	49.6	3.9	400	2 AAZ19134	Aaz19134 M. tuberc
16	49.6	3.9	400	2 AAZ19134	Aaz19134 M. tuberc
17	49.6	3.9	14800	6 ABL66291	Ab166291 Lung canc
18	49.6	3.9	14835	6 AAS94858	Aas94858 Human DNA
19	49	3.8	7175	2 AAG84658	Aag84658 Human neu
20	49	3.8	7175	2 AAV42686	Aav42686 DNA encod
21	49	3.8	7175	3 AAA71704	Aaa71704 Human cal
22	49	3.8	7175	6 AAD39956	Aad39956 Human cal
23	49	3.8	7177	7 ABZ58367	Abz58367 Human N-t

24	49	3.8	7266	2 AAV29059	Aav29059 Human cal
25	49	3.8	7362	2 AAQ37817	Aaq37817 Sequence
26	49	3.8	7362	2 AAQ84657	Aaq84657 Human neu
27	49	3.8	7362	2 AAV42685	Aav42685 DNA encod
28	49	3.8	7362	3 AAA71703	Aaa71703 Human cal
29	49	3.8	7362	6 AAD39955	Aad39955 Human cal
30	49	3.8	7364	6 ABL65869	Ab165869 Lung canc
31	49	3.8	7364	7 ABZ58366	Abz58366 Human N-t
32	49	3.8	7376	2 AAX88001	Aax88001 N-type ca
33	49	3.8	109519	5 AAS08693	Aas08693 Micromono
34	48.2	3.8	524	4 AAK92108	Aak92108 Human cDN
35	48.2	3.8	524	4 AAK93547	Aak93547 Human cDN
36	48.2	3.8	1769	4 AAK94827	Aak94827 Human ful
37	48	3.8	110000	4 AAI99683	Continuation (40 o
38	47.6	3.7	5944	6 ABL32253	Ab132253 Human imm
39	47.6	3.7	5944	6 ABL92203	Ab192203 Chemicall
40	46.8	3.7	536	9 ADB68842	Adb68842 Minority
41	46.8	3.7	1383	7 ADA71056	Ada71056 Rice gene
42	46.6	3.6	745	2 AAQ62837	Aaq62837 GGF gene
43	46.6	3.6	745	2 AAQ58330	Aaq58330 GGF gene
44	46.6	3.6	745	2 AAT30988	Aat30988 Human gli
45	46.6	3.6	745	2 AAT48077	Aat48077 Human gli

ALIGNMENTS

RESULT 1

AAD00294
ID AAD00294 standard; cDNA; 1276 BP.
XX
AC AAD00294;
XX
DT 05-SEP-2000 (first entry)
XX
DE Mouse oocyte-specific O1-180 cDNA clone.
XX
KW Oocyte-specific; ovary; O1-180; mouse; gynaecological; treatment; screen;
KW cell proliferative disorder; cell degenerative disorder; contraceptive;
KW modulator; signalling pathway; human infertility; cancer; ovulation; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 28..1113
FT /*tag= a /product= "Mouse oocyte-specific protein, O1-180"
XX
PN WO200024755-A1.
XX
PD 04-MAY-2000.
XX
PF 28-OCT-1999; 99WO-US025209.
XX
PR 28-OCT-1998; 98US-0106020P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Matzuk MM, Wang P;
XX
DR WPI; 2000-350684/30.
XX
DR P-PSDB; AAY70948.
XX
PT O1-180, O1-184 and O1-236 polypeptides and nucleic acids encoding them,
XX useful for evaluating potential contraceptives to block ovulation in a
XX reversible manner.
XX
PS Claim 2; Fig 1; 54pp; English.
XX
CC The present sequence is the cDNA encoding the mouse oocyte-specific
XX protein O1-180, expressed in the oocytes of primary (one-layer) prenatral
XX follicles through ovulation. It provides in vitro and in vivo reagents
XX for studying ovarian development and function. This sequence has

CC gynaecological and contraceptive activity. Agents which modulate O1-180,
CC O1-184 and O1-236 may be used to treat cell proliferative or degenerative
CC disorders, associated with abnormal expression of these ovary specific
CC genes. This ovary-specific sequence can be used as reagents to evaluate
CC potential contraceptives, to block ovulation in a reversible manner. It
CC is also used to screen for genetic mutations in signalling pathways, that
CC are associated with some forms of human infertility or gynaecological
CC cancers

XX SQ Sequence 1276 BP; 279 A; 389 C; 394 G; 214 T; 0 U; 0 Other;

Query Match 99.1%; Score 1265; DB 3; Length 1276;
Best Local Similarity 99.9%; Pred. No. 9.2e-291;
Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AAGCGGGGAGCGCGGAGCAGCACCATTGTTCCGGCGAGCAGTTCACCCCTGCCG 60
DB 1 AAGCGGGGAGCGCGGAGCAGCACCATTGTTCCGGCGAGCAGTTCACCCCTGCCG 60
QY 61 CATCTTATCCGAGGCCACCAAGCCGGGATGGCTGGAGGTTCGAGCCAGGGGCTGC 120
DB 61 CATCTTATCCGAGGCCACCAAGCCGGGATGGCTGGAGGTTCGAGCCAGGGGCTGC 120
QY 121 GACCCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTC 180
DB 121 GACCCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTC 180
QY 181 GACAGCACACAGGGGCAAGCTCATGGCCCTGCTGTGCGGATGGTCCCGGTCGTC 240
DB 181 GACAGCACACAGGGGCAAGCTCATGGCCCTGCTGTGCGGATGGTCCCGGTCGTC 240
QY 241 AGCAGCGTGAGCTGCGGTGAGGTGAACCCGCGCGGAGTGAACCCGCGCGGAGTGA 300
DB 241 AGCAGCGTGAGCTGCGGTGAGGTGAACCCGCGCGGAGTGAACCCGCGCGGAGTGA 300
QY 301 CTCGGCGCGCGACGCTGAGCGTGCAGGGTGCAGCGCGCGGAGCGCCCGGATCGGT 360
DB 301 CTCGGCGCGCGACGCTGAGCGTGCAGGGTGCAGCGCGCGGAGCGCCCGGATCGGT 360
QY 361 TCTGTCAACCCGTTGGGCGACGCGCGCGCGGAGATCCCGGATCTGGGAGACCGTA 420
DB 361 TCTGTCAACCCGTTGGGCGACGCGCGCGCGGAGATCCCGGATCTGGGAGACCGTA 420
QY 421 GCCCGTCTCTGCTGCGTACCTCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 480
DB 421 GCCCGTCTCTGCTGCGTACCTCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 481 AGCGAGACACCCAGGAGGAGGAGGAGCGCGGATCTCTGGGAGCGCGGAGACCGGAG 540
DB 481 AGCGAGACACCCAGGAGGAGGAGGAGCGCGGATCTCTGGGAGCGCGGAGACCGGAG 540
QY 541 CCGAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 600
DB 541 CCGAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 600
QY 601 CAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 CAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 GCGGCGATGAGTCTGAGCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 661 GCGGCGATGAGTCTGAGCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 GACCCGCGTATGAGTGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 780
DB 721 GACCCGCGTATGAGTGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 780
QY 781 AAGGAGCGCTGCGTCTTTCAGTTCTTAGAGCAGAGTACGGCTACTATCACTCAAGGAC 840
DB 781 AAGGAGCGCTGCGTCTTTCAGTTCTTAGAGCAGAGTACGGCTACTATCACTCAAGGAC 840
QY 841 TGCATAATCCGTTGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 841 TGCATAATCCGTTGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT

DB 841 TGCATAATCCGTTGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899
QY 901 CTTGAAAAGTCTTCCGAGTGTGTGAGAAATCCCTACAACTTACAGAGTGGAGACAT 960
DB 900 CTTGAAAAGTCTTCCGAGTGTGTGAGAAATCCCTACAACTTACAGAGTGGAGACAT 959
QY 961 CACCTGTCAAAAGTGTGAAAAGAACTAGATGTGCTGCGCCAGTCAGATTTTCGCCAGTGA 1020
DB 960 CACCTGTCAAAAGTGTGAAAAGAACTAGATGTGCTGCGCCAGTCAGATTTTCGCCAGTGA 1019
QY 1021 CCCTAAACGCCCCCATCGCAAGACTTGTGGGAGATGCAAGGACAAACGCCCTGTCTG 1080
DB 1020 CCCTAAACGCCCCCATCGCAAGACTTGTGGGAGATGCAAGGACAAACGCCCTGTCTG 1079
QY 1081 CGACAGCACCTTCAGCTTCAATATCATATTTAGTGAGAGTCGAAAACGTTTCTGCTAGA 1140
DB 1080 CGACAGCACCTTCAGCTTCAATATCATATTTAGTGAGAGTCGAAAACGTTTCTGCTAGA 1139
QY 1141 TGGGGCTTAATGGAATGGACAAGTGAGCTTTCTCCCTCTTTCACCTCTTCCCTTTCCAAAT 1200
DB 1140 TGGGGCTTAATGGAATGGACAAGTGAGCTTTCTCCCTCTTTCACCTCTTCCCTTTCCAAAT 1199
QY 1201 TCTTCATGACAGACAGTGTACTTGGATATAAAGCTGTGTAATAAAGTATTTGCAACA 1260
DB 1200 TCTTCATGACAGACAGTGTACTTGGATATAAAGCTGTGTAATAAAGTATTTGCAACA 1259
QY 1261 AAAAAAATAAAAAA 1277
DB 1260 AAAAAAATAAAAAA 1276

RESULT 2
ABZ24589
ID ABZ24589 standard; cDNA; 1258 BP.
XX
AC ABZ24589;
XX
DT 31-MAR-2003 (first entry)
XX
DE Mouse ovary-specific O1-180 cDNA.
XX
KW Ovary; O1-180; mouse; contraceptive; antiinfertility; cytostatic;
KW gene therapy; gene; ss.
OS Mus musculus.
FH Key Location/Qualifiers
CDS 26..1108
FT /*tag= a
FT /product= "Murine O1-180"
FT /transl_except= (pos:182..184,aa:Ser)
FT /transl_except= (pos:1004..1007,aa:Phe)
XX
PN WO200288314-A2.
XX
PD 07-NOV-2002.
XX
PF 26-APR-2002; 2002WO-US013245.
XX
PR 27-APR-2001; 2001US-00844864.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PA (AMHP) WYETH.
XX
XX Matzuk MM, Wang P, Bai Y, Wu X;
XX
XX WPI; 2003-167110/16.
XX P-PSDB; ABP58235.
XX
XX New ovary-specific-genes comprising O1-180 or O1-236, useful for
XX decreasing conception or enhancing fertility, or for the preparation of a
XX composition for treating e.g. cancer.

PS Claim 2; Page 136-137; 141pp; English.

XX The present sequence is that of a human ovary-specific OI-180
CC polynucleotide. In mice, loss of OI-180 is associated with female
CC infertility and subfertility. The invention provides ovary-specific and
CC oocyte-specific murine and human OI-180, OI-184 and OI-236
CC polynucleotides and polypeptides. These genes and their protein products
CC appear to relate to various cell proliferative or degenerative disorders,
CC especially those involving ovarian tumours, such as germ line tumours and
CC granulosa cell tumours, or infertility, such as premature ovarian
CC failure. The invention provides a method for detection of a cell
CC with the expression of OI-180, OI-184 or OI-236. It also provides a
CC method for treating such disorders by using an agent which suppresses or
CC enhances the respective activities of OI-180, OI-184 or OI-236, and a
CC method of screening for compounds that interact and/or modulate the
CC expression or activity of the ovary-specific genes. These compounds are
CC possible contraceptive or fertility enhancing agents. The modulator is
CC preferably a polypeptide, small molecule or polynucleotide sequence

XX Sequence 2075 BP; 588 A; 413 C; 473 G; 601 T; 0 U; 0 Other;

Query Match 10.5%; Score 134.6; DB 7; Length 2075;
Best Local Similarity 71.1%; Pred. No. 6.6e-22;
Matches 207; Conservative 0; Mismatches 79; Indels 5; Gaps 2;
QY 968 CAAAGTTGTAAGAACTAGATGTCCTGCCAGTCAGATTTCGCCAGTGGACCTTAA 1027
DB 1360 CAGAGTTGTAACAACAGAGATGTTCTCTGCCAGTAAACTTCGCCAGTGGACCTTAA 1419
QY 1028 CCCCCCATCGCAAGACTTGTGGAGATCAAGGCAAAAGCGCTGTCTCGCACAGC 1087
DB 1420 CGGCCCCACCGTCAAGATTGTGGGTAGATCAAGGCAAAAGCGCTGTCTGTGACAGC 1479
QY 1088 ACCTTCAGCTTCAATACATCAATTATTA-GTGAAGTCGAAAGAGTTTCTCTAGATGGGC 1146
DB 1480 ACTTTCAGCTTCAATATACATCATTTAGTGAAGTCAAGTGTGTGTGATCGCTGATG 1539
QY 1147 TAATGAATGGACAGTGAAGCTTCTCCCTTCTCACTCTTCCCTTCCAAATTTCTCA 1206
DB 1540 GAGTAGACAGTGAAGCTTTTTCGCTGCTCTCTCACTCTTCCCTTCTCAAAATTTCTCA 1599
QY 1207 TGACAGACAGTGTACTTCGATATATAAGCTGTGATATAAGAGTATTGCAA 1257
DB 1600 TGAAGGCAAGTATTCTG-----AAAAGCCTTCAATAAAGTATTGCAA 1646

RESULT 6

ID AAX53491 standard; DNA; 114955 BP.

XX AAX53491;

AC AAX53491;

DT 05-JUL-1999 (first entry)

XX Human adenosine A1 receptor antisense oligonucleotide fragment.

XX Antisense oligonucleotide; multiple target; antisense treatment;
KW impaired respiration; inflammation; lung disease;
KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
KW acute asthma; allergy; asthma; impeded respiration;
KW respiratory distress syndrome; pain; cystic fibrosis;
KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
KW prostate cancer; ss.

XX Synthetic.

OS WO9913886-A1.

XX 25-MAR-1999.

PD

XX 17-SEP-1998; 98WO-US019419.
XX 17-SEP-1997; 97US-0059160P.
PR 09-JUN-1998; 98US-00093972.
XX (UYEC-) UNIV EAST CAROLINA.
XX Nyce JW;
XX WPI; 1999-229400/19.
DR New antisense oligonucleotides used in treatment of, e.g. pulmonary
XX vasoconstriction.
XX Disclosure; Page 37; 120pp; English.
XX The specification describes antisense oligonucleotides (AAX52869-X55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, gene initiation
CC codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-
CC end and the juxta-section between coding and non-coding regions and all
CC segments of RNAs encoding proteins associated with one or more diseases,
CC from sequences AAX55272-74. These multiple target oligonucleotides
CC (specifically AAX55180-271) can be used for the antisense treatment of
CC diseases and conditions. Typical diseases and conditions are those
CC associated with impaired respiration and inflammation, including lung
CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,
CC acute asthma, allergies, asthma, impeded respiration, respiratory
CC distress syndrome, pain, cystic fibrosis, pulmonary hypertension,
CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary
CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.
CC colon cancer, breast cancer, lung cancer, pancreatic cancer,
CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as
CC well as all types of cancers which may metastasize or have metastasized
CC to the lungs, including breast and prostate cancer
XX SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;
Query Match 4.6%; Score 58.8; DB 2; Length 114955;
Best Local Similarity 32.6%; Pred. No. 0.0025;
Matches 251; Conservative 71; Mismatches 441; Indels 8; Gaps 3;
QY 3 GCGCGCGGAGCGCGGAGCGACCCATGTTCCCGCGAGACGTTCCACCCCTGCCGCA 62
DB 104641 GCGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104700
QY 63 TCCTTATCGGAGGCGACCAAGCGGCGGATGGTGGAGTTCGGAGTTCGGAGCGGCGGCGG 122
DB 104701 GC-----GCGCSNNNDNNCCGCGGCGGCGGCGGCGGCGGCGGCGGCSNN 104756
QY 123 ACCCGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
DB 104757 NNNDNNCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104816
QY 183 CAGCCACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
DB 104817 SNNNDNNCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104876
QY 243 CAGCGGTGACGCTGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 104877 NCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104936
QY 301 CTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 104937 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104996
QY 361 TCCTGTCAAACCGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 104997 GSNNDNNBGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 105056
QY 421 GCGCGGCTTCTGCTCGGTGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

Mon Sep 27 09:07:51 2004

```
XX 17-FEB-1999; 99WO-US003265.
XX
XX 18-FEB-1998; 98US-00024753.
XX
XX 05-MAY-1998; 98US-00072596.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedwick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
XX WPI; 1999-527416/44.
XX
XX New polypeptide comprising antigenic portions of M. tuberculosis.
XX
XX Claim 4; Page 215-216; 323pp; English.
XX
XX This invention describes novel recombinant antigens and their encoding
XX nucleic acids derived from Mycobacterium tuberculosis. The novel
XX polypeptides are useful for detecting M. tuberculosis infection in a
XX biological sample by detecting antibodies which bind with the
XX polypeptides, and are useful as vaccines for immunizing against M.
XX tuberculosis infection. The new detection methods are needed as current
XX vaccination strategies do not provide 100% immunity
XX
XX Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 U; 0 Other;
XX
Query Match 3.9%; Score 49.6; DB 2; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.064;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 458 CCTCACTGGAGGTTCGGGAGGCGAGCAGACACCCACGAGGGAGGGAGGCCGGCAT 517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 CTTTACCCCAAGGCGCGGCAACCGCGGCAACCGCGGCGTACGCGGGGTGGCGGCAA 178
QY 518 CCTCGGGGACCGGGAAACCGAGAGGTGGCGCGAGGAAAGCGGTCCCCCAGC 577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 CGCGGAAACGGCGGAAACGGCGGAGACACACCACCGCGCGCGGACACAGG 238
QY 578 CGCGAAGCGAGAGGGGCGATGTTTCAGGTCAGGGGAGCGCGGTTGGAGCAGCCAC 637
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 CGGCGACGGCGGGCGCGGGCGCGGCGGAAACCGCGGAAACCGCGGAGCGCGGCAC 298
QY 638 CACCGGAGGACCGGAACAGTGTGGCGCGATGCTGAGCTGGAGCGAGGAGCCAT 697
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 CGGCACCGCGGCGCAACAGGCAACCGCGGCAACCGCGGCGCACCGCGGCAAGCGGCAC 358
QY 698 GTCTGCGCGAGAGATGGCTCAGG 721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 CGGCGGCGAGGTGCACTCTCAGG 382
```

Search completed: September 24, 2004, 17:34:43
Job time : 587 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 17:17:47; Search time 114 Seconds
(without alignments)
6216.424 Million cell updates/sec

Title: US-09-830-810A-1
Perfect score: 1277
Sequence: 1 aagcgggcgggcgcgggga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	4.2	5467	1	US-07-745-206A-12
2	54	4.2	5467	2	US-08-311-363-12
3	53.2	4.2	7218	1	US-08-232-463-14
4	50.8	4.0	6232	3	US-08-456-200B-11
5	49.6	3.9	400	4	US-09-056-556-179
6	49.6	3.9	400	4	US-09-072-596-174
7	49.6	3.9	400	4	US-09-072-967-179
8	49	3.8	7175	1	US-08-455-543A-8
9	49	3.8	7175	2	US-08-193-078B-8
10	49	3.8	7175	2	US-08-223-305C-8
11	49	3.8	7175	3	US-08-149-097D-8
12	49	3.8	7175	3	US-08-949-386-8
13	49	3.8	7175	3	US-08-450-562-8
14	49	3.8	7175	4	US-08-984-709A-8
15	49	3.8	7175	4	US-08-450-272-8
16	49	3.8	7175	4	US-08-450-273-8
17	49	3.8	7177	4	US-09-268-163-7
18	49	3.8	7266	3	US-08-713-118-1
19	49	3.8	7266	3	US-09-452-007-1
20	49	3.8	7362	1	US-08-455-543A-7
21	49	3.8	7362	2	US-08-193-078B-7
22	49	3.8	7362	2	US-08-223-305C-7
23	49	3.8	7362	2	US-08-149-097D-7
24	49	3.8	7362	3	US-08-949-386-7
25	49	3.8	7362	3	US-08-450-562-7
26	49	3.8	7362	4	US-08-984-709A-7
27	49	3.8	7362	4	US-08-450-272-7

28	49	3.8	7362	4	US-08-450-273-7	Sequence 7, Appli
29	49	3.8	7362	4	US-09-268-163-5	Sequence 5, Appli
30	49	3.8	7376	4	US-09-268-163-3	Sequence 3, Appli
31	48	3.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 32	46.6	3.6	745	1	US-08-036-555B-163	Sequence 163, App
C 33	46.6	3.6	745	1	US-08-469-569-163	Sequence 163, App
C 34	46.6	3.6	745	1	US-08-249-322A-163	Sequence 163, App
C 35	46.6	3.6	745	1	US-08-469-528A-163	Sequence 163, App
C 36	46.6	3.6	745	2	US-08-734-591A-163	Sequence 163, App
C 37	46.6	3.6	745	2	US-08-469-660-163	Sequence 163, App
C 38	46.6	3.6	745	3	US-08-341-018-51	Sequence 51, Appl
C 39	46.6	3.6	745	3	US-08-470-335-163	Sequence 163, App
C 40	46.6	3.6	745	3	US-08-735-021-163	Sequence 163, App
C 41	46.6	3.6	745	3	US-08-734-664A-163	Sequence 163, App
C 42	46.6	3.6	745	3	US-08-470-339-163	Sequence 163, App
C 43	46.6	3.6	745	4	US-08-467-603-163	Sequence 163, App
C 44	46.6	3.6	745	5	PCT-US94-05083C-159	Sequence 159, App
C 45	46.6	3.6	745	5	PCT-US95-06846A-163	Sequence 163, App

ALIGNMENTS

RESULT 1
US-07-745-206A-12
; Sequence 12, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5467 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
; LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585
; LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744
; LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010
; LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGTGCAGGTGAACCCGCGCGGACGCTCGTGCAGTGTCTACTCGGGCC 309
Db 2751 GAGGCCCGAAGCGAGAGCGGGAGCCCGTGCCTCGGAGAGCGCGCGCGCAC 2810
QY 310 GCGAGCTGCGCTGCGAGGTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 369
Db 2811 GCAC-GCCACAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2867
QY 370 CCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 429
Db 2868 GCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCCACGAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 549
Db 2988 GCGCCAAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGCGAGGAAAGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 609
Db 3048 GCGGAGAGCGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GGGCAGCGCGGTGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 668
Db 3108 CACGAGGCTGTGAGAAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCG 3166

RESULT 9
US-08-193-078B-8
; Sequence 8, Application US/08193078B
; Patent No. 5846757
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
; STREET: 1660 UNION STREET
; CITY: SAN DIEGO
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,078B
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999

TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
; US-08-193-078B-8
Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGTGCAGGTGAACCCGCGCGGACGCTCGTGCAGTGTCTACTCGGGCC 309
Db 2751 GAGGCCCGAAGCGAGAGCGGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2810
QY 310 GCGAGCTGCGCTGCGAGGTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 369
Db 2811 GCAC-GCCACAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2867
QY 370 CCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 429
Db 2868 GCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCCACGAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 549
Db 2988 GCGCCAAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGCGAGGAAAGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 609
Db 3048 GCGGAGAGCGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GGGCAGCGCGGTGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 668
Db 3108 CACGAGGCTGTGAGAAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCG 3166

RESULT 10
US-08-223-305C-8
; Sequence 8, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926

FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-55038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-149-097D-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GAGCGTGGGTGAGTGAACCGCGCGCGAGCGCTCGGTGAGTGTTCACCTCGGGCGC 309
Db 2751 GAGGCCCCGAGCGGAGAGCGGGAGCGCGGTGCGGGAGAGCGCGCGCGCGCAC 2810
QY 310 CGCAGCGCTGAGCTGCGAGTGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 369
Db 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2867
QY 370 CCGCGTGGCCACCGCGCGCGCGCGAGATCCCGCGCATCTCTGGCAGACCGTAGCCCGTTC 429
Db 2868 GGCCTCAGGCCCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCGCAGAGGAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GGGGCCAAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047
QY 550 GTGCCCGCGAGGAAAGCGGTCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609

Query Match 3.8%; Score 49; DB 3; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GAGCGTGGGTGAGTGAACCGCGCGCGAGCGCTCGGTGAGTGTTCACCTCGGGCGC 309
Db 2751 GAGGCCCCGAGCGGAGAGCGGGAGCGCGGTGCGGGAGAGCGCGCGCGCGCGCAC 2810
QY 310 CGCAGCGCTGAGCTGCGAGTGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 369
Db 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2867
QY 370 CCGCGTGGCCACCGCGCGCGCGCGAGATCCCGCGCATCTCTGGCAGACCGTAGCCCGTTC 429
Db 2868 GGCCTCAGGCCCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCGCAGAGGAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GGGGCCAAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047
QY 550 GTGCCCGCGAGGAAAGCGGTCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609

US-08-149-097D-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GAGCGTGGGTGAGTGAACCGCGCGCGAGCGCTCGGTGAGTGTTCACCTCGGGCGC 309
Db 2751 GAGGCCCCGAGCGGAGAGCGGGAGCGCGGTGCGGGAGAGCGCGCGCGCGCAC 2810
QY 310 CGCAGCGCTGAGCTGCGAGTGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 369
Db 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2867
QY 370 CCGCGTGGCCACCGCGCGCGCGAGATCCCGCGCATCTCTGGCAGACCGTAGCCCGTTC 429
Db 2868 GGCCTCAGGCCCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCGCAGAGGAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GGGGCCAAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047
QY 610 GGGCAGCGCGGTGGGAGCAGCAGCACCACCGCGAGGAGCGCGCGCGCGCGCGCGAT 668
Db 3108 CACGAGCTGTGAGAGAGGAGACCAACCGAGAGGAGGAGCGCGCGCGCGCGCGAT 3166

RESULT 12
US-08-949-386-8
Sequence 8, Application US/08949386
Patent No. 6090623
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McQue, Ann
APPLICANT: Gillespie, Allison
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:

3108 CACGAGGCTGTGGAGAGGAGGACCCACGAGAGGAGGCCACCGAGAGGAGGCTGAGAT 3166

RESULT 14
US-08-984-709A-8
Sequence 8, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Staederman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, Suite 700
City: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-984-709A-8

Query Match 3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps

QY 250 GACGCTGCGGTGCAGGTGAACCCCGCGCGGAGCGCTCGGTGACGTGTTCACTCGGGCGC 3166
DB 2751 GAGGCCCCGAAGGCGGAGAGCGGGGAGCGCGGTGTCGCCGGAGGAGCGCGCGCGCGCAC 2888

QY 310 CGCAGCTGCAGCCTGCAGAGGTGCGGAGCAGCCCGACGCCCGATCGGTTCTCTGTCAA 3666
DB 2811 CGCA-GCCACAGCAAGAGAGGCCGCGGGCC--CCCGAGGCGCGGAGCGAGCGCGCCGA 2888

QY 370 CCCCGTGGCCACGCGCGCGCGCGGAGATGCCCGCGATCTCTGGCAGACCGGTAGCCCCGTT 4200
DB 2868 GSCCCAGGCCCCGAGGGCGCGCGCGGACACACCGCGCGCGCTCCCCGAGAGGCGGCC 2933

Search completed: September 24, 2004, 20:03:21
Job time : 121 secs

Blank